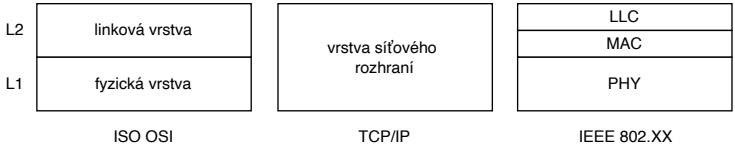
Počítačové sítě 1

linková vrstva

Martin Trnečka

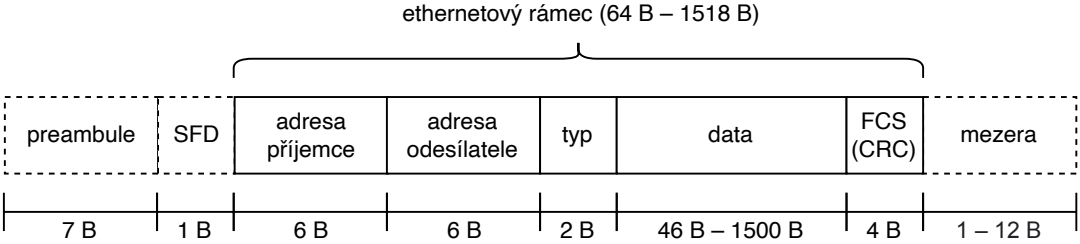
Katedra informatiky
Univerzita Palackého v Olomouci

TCP/IP architektura



- fyzická vrstva přenáší bity
- na úrovni linkové vrstvy přenosová jednotka: *linkový rámec*
- podobu určuje daná technologie

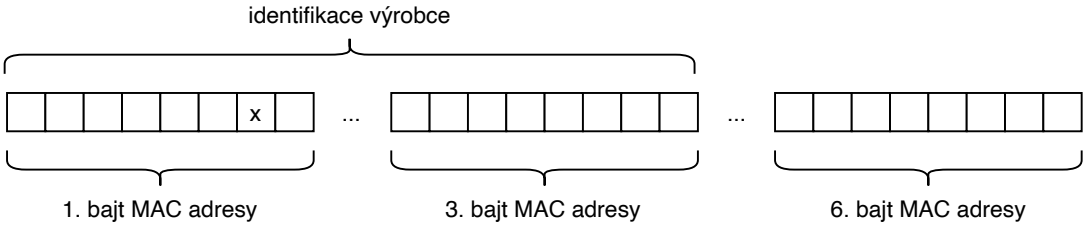
Ethernetový rámec (Ethernet II)



- konkrétní příklad linkového rámce
  - alternativa IEEE 802.3 (nutný SNAP)
  - rozšíření VLAN (IEEE 802.1q)
  - Jumbo Frame
- součást fyzické vrstvy
  - preambule – synchronizace, tvar 10101010
  - SFD (start frame delimiter) – signalizace začátku rámce, tvar 10101011
  - mezera – pauze ve vysílání (velikost dle použité technologie)

MAC adresa

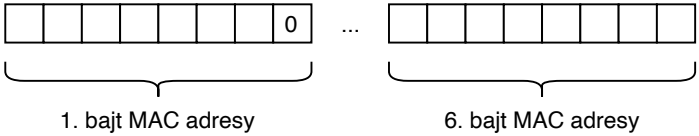
- adresa odesílatele a příjemce
- adresa fyzického rozhraní (network interface card, NIC)
- 6 B, zapisované hexadecimálními číslicemi
- tvar: `XX:XX:XX:XX:XX:XX`, `X` je 0–9, a–e
- například: `38:f9:d3:0b:ed:3f`



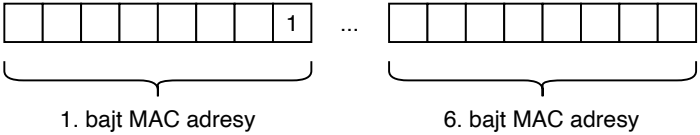
- `x = 0` jedinečná globální adresa, `x = 1` vlastní adresace
- bity jednotlivých bajtů jsou posílány v opačném pořadí než jsou zapsány (kvůli speciálním MAC adresám)

Speciální MAC adresy

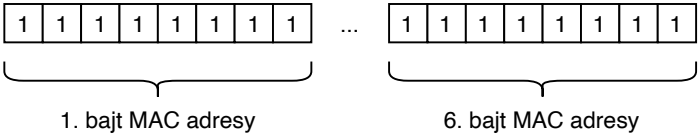
- unicast – nejméně významný bit prvního bajtu je 0



- multicast – nejméně významný bit prvního bajtu je 1



- broadcast – všechny bity jsou 1 → `ff:ff:ff:ff:ff:ff`

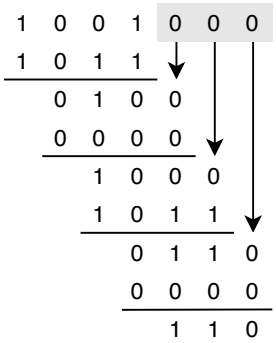


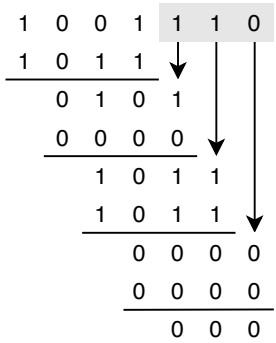
Ethernetový rámec

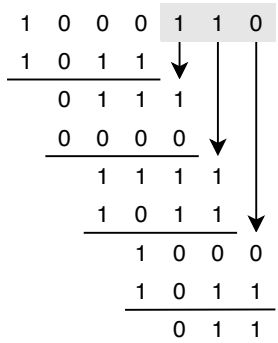
- položka *typ* identifikuje protokol vyšší vrstvy
  - 0–1500 vyhrazeno pro IEEE 802.3 (délka dat)
  - 2054 (desítkově) protokol ARP
  - 2048 (desítkově) IPv4
- položka *data* data z vyšší vrstvy
  - 46 B – 1500 B
  - pokud je dat méně je třeba doplnit na velikost 46 B (detekce kolizí)
  - v případě IEEE 802.3 navíc hlavička SNAP $\rightarrow$ lze přenést méně dat
- FSC – detekce chyb (Cyclic Redundancy Check, CRC)
  - pouze kontrola správnosti
  - nesprávné rámce jsou zahozeny
  - absence potvrzování $\rightarrow$ nespolehlivý přenos
  - při přenosech s vysokou spolehlivostí zbytečná zátěž

Odbočka: Idea výpočtu CRC

- data = 1001, generátor = 1011, data se rozšíří o n nul, kde n počet bitů generátoru - 1, rozšířená data se vydělí v modulo 2 aritmetice (operace XOR), zbytek po dělení je použit jako CRC
- kontrola = zbytek po dělení musí být nulový, jinak chyba







Odbočka: Režie protokolu

- řídící informace (hlavička) = režie
- na úkor přenášených dat

$$\text{efektivita protokolu} = \frac{\text{velikost dat}}{\text{velikost rámce}}$$

- maximální efektivita = $\frac{1500}{1518} = 0{,}99$
- minimální efektivita = $\frac{64}{82} = 0{,}78$

Propojování sítí

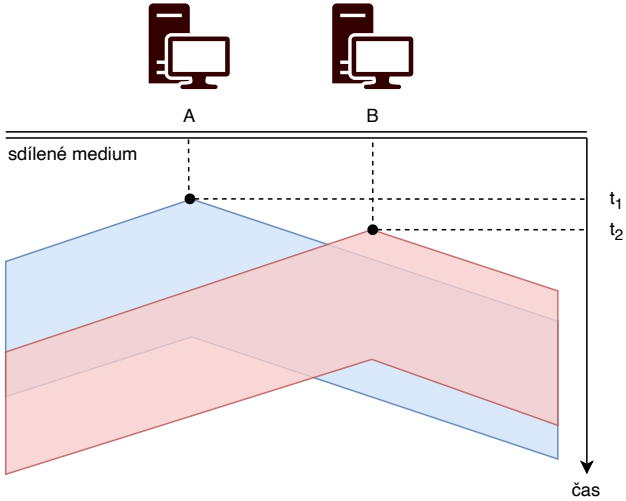
- oddělení kolizních domén (síťových segmentů)
- různé fyzické topologie
  - sběrnicová
  - kruhová
  - hvězdicová
  - hybridní (hvězdicovo-kruhová, hvězdicovo-sběrnicová)
  - mesh (full, partial)
- sdílené médium $\rightarrow$ řízení přístupu

Kolize při přístupu ke sdílenému médiu

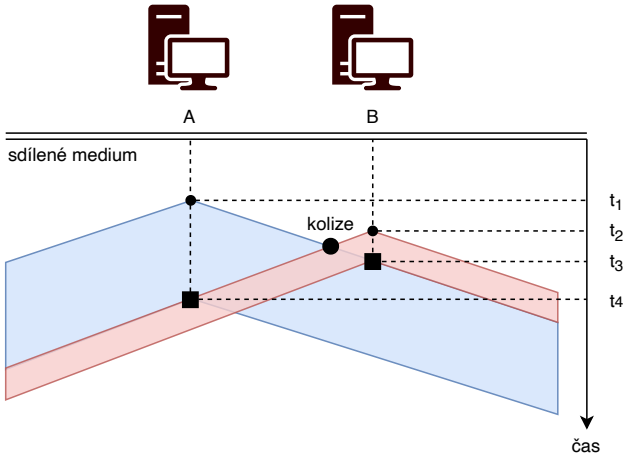
- kolize = smíchání signálů
- více uzlů vysílá současně → znehodnocení signálu
- detekce porovnáním vysílaného a přijímaného signálu
- protokol CSMA (Carrier Sense Multiple Access)
  - CSMA → „sense before submit“
  - CSMA/CD (Collision Detection) → detekce kolize
  - CSMA/CA (Collision Avoidance) → předcházení kolizi (bezdrátové sítě)

Kolize v CSMA

- stále může dojít ke kolizi (v důsledku zpoždění šíření)



Detekce kolize v CSMA/CD



Detekce kolize v CSMA/CD

- postup
  1. uzel poslouchá zda je médium volné $\rightarrow$ začne vysílat
  2. uzel poslouchá zda nevysílá jiný uzel $\rightarrow$ přeruší vysílání a vyšle JAM
  3. stanice se na náhodný čas odmlčí (odvozen z MAC, několik pokusů, při každém se zdvojnásobí)
- detekce kolize funguje jen v době vysílání $\rightarrow$ rámec musí být vysílán $2\times$ déle než je doba zpoždění $\rightarrow$ minimální délka rámce
- například: maximální zpoždění v ethernetu je 25,6 µs $\rightarrow$ je třeba vysílat 51,2 µs při rychlosti 10 Mb/s $\rightarrow$ 512 bitů = 64 B
- více uzlů, větší vzdáleností, větší provoz $\rightarrow$ více kolizí
- moderní ethernet CSMA/CD nevyužívá (nedochází ke kolizím, plně duplexní, přepínaný provoz)

Řízení toku dat na úrovni L2

- různé rychlosti zpracování na straně odesílatele a příjemce $\rightarrow$ zahlcení příjemce $=$ ztráta dat
- přenos v prostředí bez šumu
  - simple protocol (žádné řízení)
  - stop and wait (po odeslání se čeká na potvrzení)
- přenos v prostředí se šumem
  - stop and wait s opakováním
  - sliding window (různé varianty, potvrzení několika rámců současně)
- v TCP/IP se nevyužívá $\rightarrow$ přenos na úrovni linkové vrstvy je nespolehlivý

Propojování sítí

- propojovací zařízení
  - opakovač (L1)
  - hub (L1)
  - switch (L2, L3)
  - router (L3)
  - další zařízení (L4, L4–7)
  - zatím se omezíme na vrstvu L2

Switch

- propojení různých LAN (obecně i různé fyzické technologie) = *síťové segmenty*
- na úrovni L2 filtrování a řízení na základě MAC adresy
  - filtrační tabulka (MAC adresa, port)
  - manuální nebo automatické plnění

Switch vs Bridge

- bridge je omezený předchůdce switche
- zajišťuje propojení na úrovni L2
- dnes se již nepoužívá, pojem *network bridge* vyhrazen pro propojení na úrovni L2
- zejména v IEEE standardech
- bridge je realizován pomocí switche (switch realizuje bridge)
- transparentní bridge (IEEE 802.1d)
  - uzly netuší, že bridge existuje
  - automatické učení
  - třeba zabránit smyčkám v topologii → cyklický oběh rámců

Automatické plnění

- uloží do tabulky:
  - adresu odesílatele
  - port kterým rámec přišel
  - čas
- pokud není adresa příjemce v tabulce $\rightarrow$ rámec je opakován na všechny porty, kromě portu ze kterého byl rámec přijat
- pokud je adresa příjemce v tabulce, příslušný port je různý od portu, kterým by rámec přijat a port není blokován $\rightarrow$ rámec je odeslán na daný port
- staré záznamy jsou po určitém čase přepsány
- poznámka: nebezpečné, částečně řeší, že NIC ignoruje rámce s jinou MAC, ale promiskuitní režim

Přepínání rámců

- rámce putují po linkách a přepínačích
- rámec o velikosti l bitů, přes linku o rychlosti r b/s $\rightarrow \frac{l}{r}$
- metoda store-and-forward, cut-through
- dočasné uložení v bufferu přepínače
- n linek $\rightarrow n - 1$ přepínačů
- přenos rámce přes n linek $\rightarrow n \cdot \frac{l}{r}$
- $\rightarrow$ zpoždění

Přepínání rámců: Zpoždění

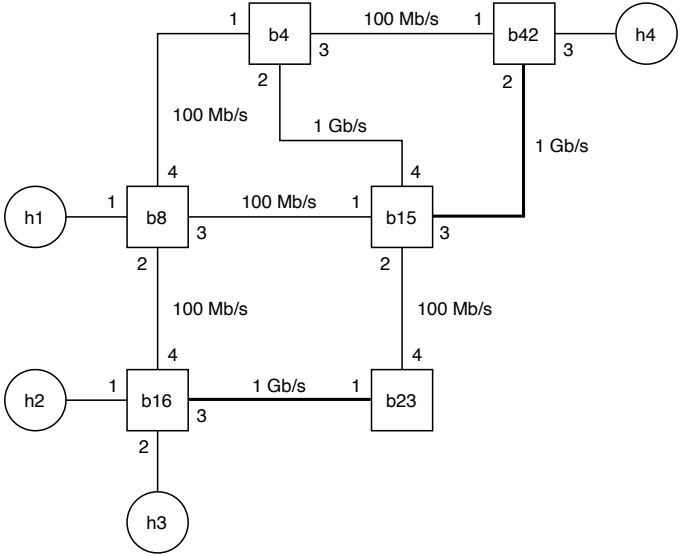
- příchozí rámce jsou ukládány do fronty a postupně odesílány
- rámce přicházejí z více zdrojů $\rightarrow$ zpoždění ve frontě
- plná fronta $=$ ztráta rámce
- zpoždění na uzlu sítě:
  - zpracování (přečtení režijní informace)
  - fronta (velmi komplexní problematika)
  - přenos
  - šíření
- propustnost sítě

Spanning-Tree Protokol (SPT)

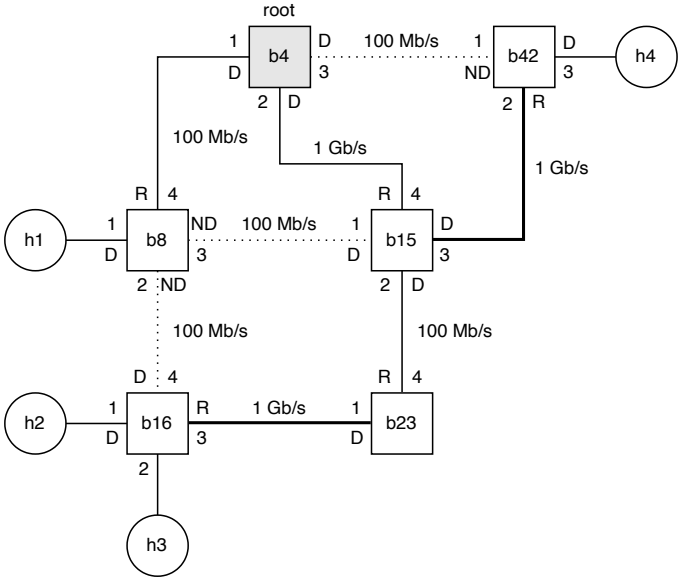
- cíl: zajistit stromovou topologii sítě (bez smyček)
- realizováno zasíláním BPDU rámce

1. vybere se *root bridge* = bridge s nejnižším id
   - id určeno MAC adresou a prioritou (lze nastavit)
   - na začátku se každý bridge označí jako root a pošle konfigurační informaci
2. vybraný root bridge nastaví všechny své porty jako *designated* (předávají rámce)
3. ostatní bridge určí své *root porty*
   - cesta s nejnižší cenou (ke kořenovému portu)
   - pokud jsou dvě cesty se stejnou cenou vybere se ta s nižším číslem portu, ostatní se vypnou *non-designated* (nepředávají rámce)
4. bridge nastaví nenastavené porty
   - na spojích s root portem na designated
   - na spojích bez root portu bridge s nižším id nastaví daný port na *non-designated*

STP: Příklad



STP: Příklad – řešení



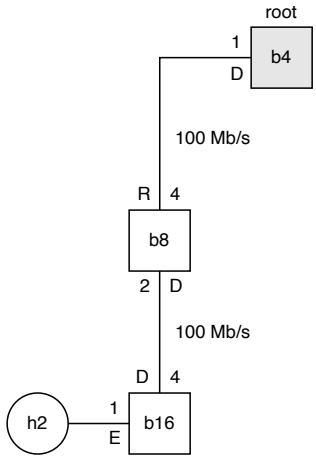
Odbočka: STP praktické poznámky

- změny v topologii $\rightarrow$ zaslání BPDU rámce
- stavy portů (světlo na portu):
  - disabled – port je neaktivní (vypnut)
  - blocking – pouze příjem BPDU (non-designated)
  - forwarding – přeposílání rámců
- root bridge zasílá Hello BPDU (každé 2s), při neobdržení
  - blocking port se po 20 s přepne na listening
  - listening – příjem BPDU rámců a jejich zpracovávání, 15 s poté learning
  - learning – učení MAC adres, rámce nejsou přeposílány 15 s, poté forwarding
- až 50 s než dojde ke změně topologie, 30 s při připojení nového uzlu
- zrychlení připojení $\rightarrow$ portfast
- vylepšení RSTP (Rapid Spanning-Tree Protocol)

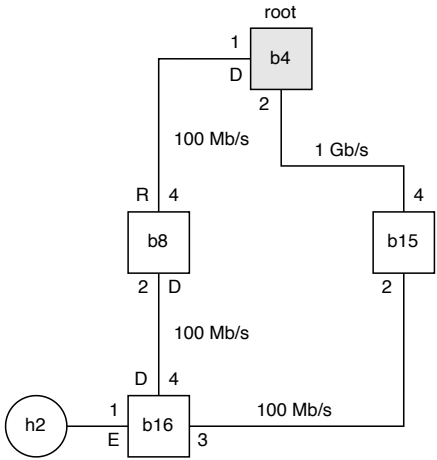
Rapid STP

- zpětná kompatibilita (při detekci RSTP a STP $\rightarrow$ STP)
- výrazně rychlejší konvergence při změně topologie
  - méně stavů: disabled, discarding, learning, forwarding
  - porty: root, designated, alternate, backup, edge
  - jiný BPDU rámec
  - nepoužívají se časovače
  - $3\times$ nepřijetí Hello BPDU $\rightarrow$ změna topologie
  - handshake (proposal-agreement)
  - je třeba zabránit smyčkám během handshake (oblokují se porty, kromě edge, dokud není handshake dokončen)
- BPDU se šíří sítí (STP nejprve je notifikován root bridge, ten notifikuje ostatní)
- STP a RSTP nevyužívají plně síťovou infrastrukturu $\rightarrow$ Shortest Path Bridging (SPB) (lze využít redundantní spoje)

RST: Příklad

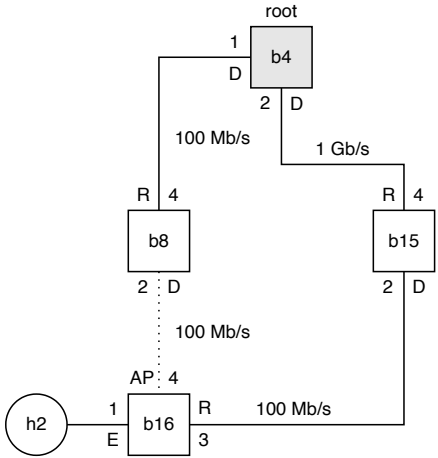


RST: Příklad



- všechny porty dotčené připojením jsou discarding (výchozí stav)
- předpokládejme, že b4 zahájí handshake s b15

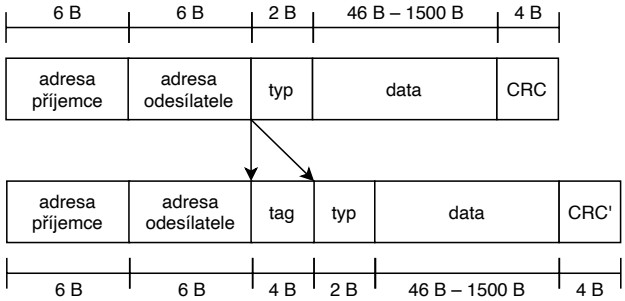
RST: Příklad – finální topologie



Virtual LAN (VLAN)

- logická síť vytvořené ve fyzické síti (IEEE 802.1q)
  - segmentace sítě
  - bezpečnost
  - omezení provozu (broatcastové domény)
- porty přepínače jsou přiřazeny do VLAN (manuálně, automaticky L3)
- identifikace `VLANX`, `X` je identifikátor 1–4095 (12 bitů)
- tagging – rozšíření záhlaví linkového rámce
- trunking – přenos tagovaných linkových rámců mezi přepínači
- access port, trunk port
- výchozí `VLAN1`, nativní (netagované rámce přes trunk)
- VLAN jsou oddělené $\rightarrow$ více způsobů řešení (router, router-on-stick, L3 switch)

Tagování


6 B
6 B
2 B
46 B – 1500 B
4 B
adresa příjemce
adresa odesílatele
typ
data
CRC
adresa příjemce
adresa odesílatele
tag
typ
data
CRC'
6 B
6 B
4 B
2 B
46 B – 1500 B
4 B

Linková vrstva

- ukázali jsem pouze Ethernet
- další technologie FDDI, Wi-Fi, Token Ring, Frame Relay, ATM
- bezdrátový přenos $\rightarrow$ jiné problémy
- další protokoly
  - Point-to-Point Protocol (PPP)
  - Point-to-Point Protocol over Ethernet (PPPoE)
  - High-Level Data Link Control (HDLC)
  - Frame Relay

Bezpečnost linkové vrstvy v LAN

- kontrolní součet $\rightarrow$ pouze chyby přenosu
  - snadný výpočet $\rightarrow$ lze měnit data v rámci
- identifikace na základě MAC adresy
  - filtrování
  - lze podvrhnout
- promiskuitní režim (dle MAC adresy)
- filtrace a blokování BPDU rámců

Odbočka: Wireshark

- nástroj pro analýzu síťové komunikace
- ukázka