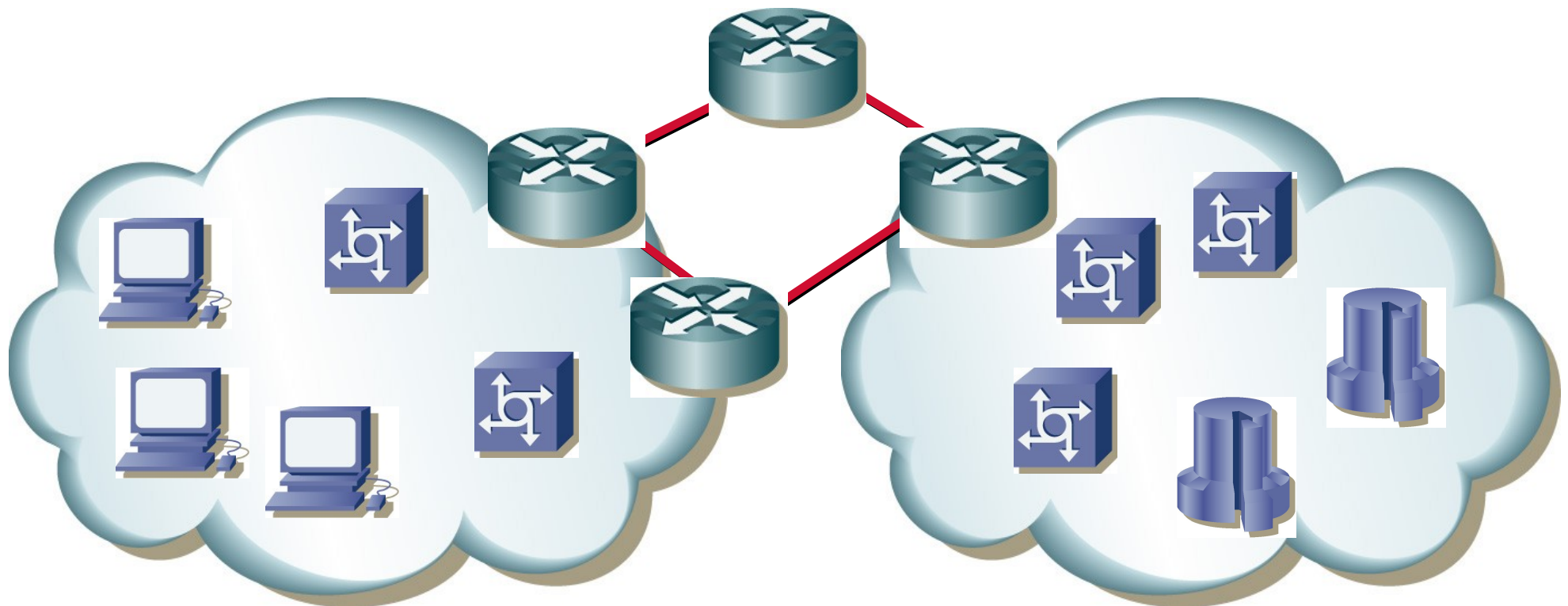


Počítačové sítě I

9. Internetworking

Miroslav Spousta, 2005

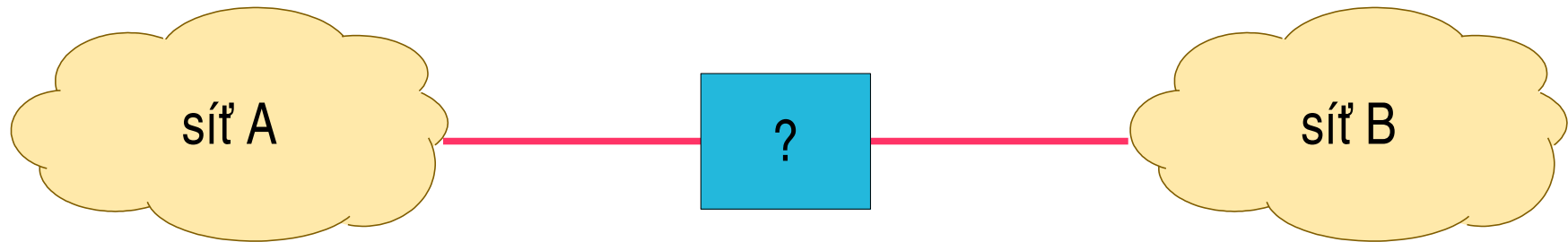
<qiq@ucw.cz>, <http://www.ucw.cz/~qiq/vsfs/>



Internetworking

- propojování sítí a jejich částí (segmentů)
- spojováním sítí vzniká inter-network
 - neboli internet
 - Internet je ta *jedna* síť
- proč propojovat sítě (a jak):
 - zvětšení možné vzdálenosti mezi uzly, zvýšení počtu stanic
 - optimalizace prostředků, provozu
 - zpřístupnění vzdálených zdrojů, prostředků
 - zlepšení služeb (např. WWW, pošta je tím užitečnější, čím více lidí ji používá)

Propojování sítí



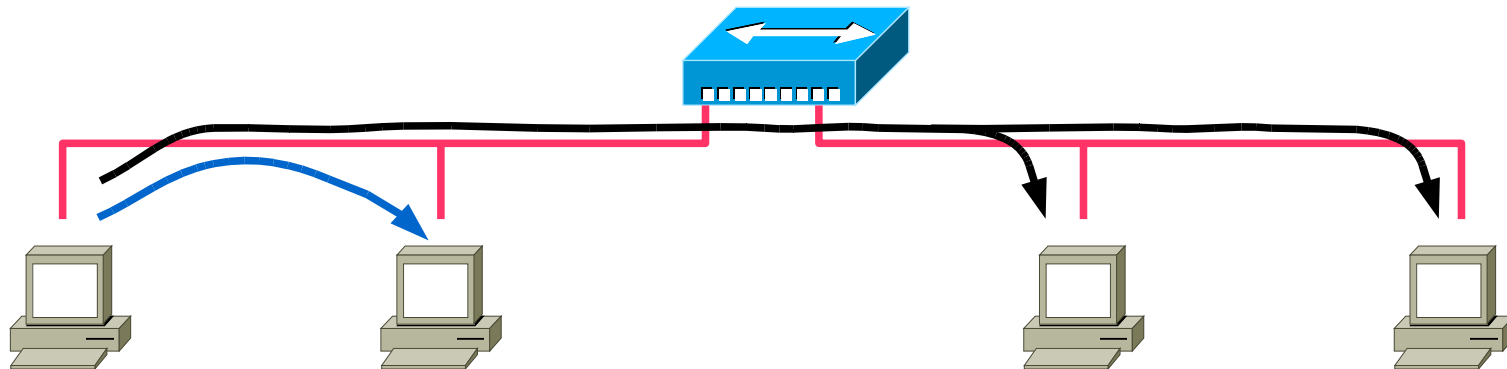
- sítě mohou být propojeny pomocí několika druhů zařízení
- na různých úrovních (vrstvách modelu OSI)
 - fyzické (opakovače)
 - linkové (switche, bridge)
 - síťové (routery)
 - aplikační (proxy, aplikační firewally)

Opakovač

- neboli repeater, propojuje zařízení na fyzické (nejnižší) vrstvě
 - někdy se označením HUB myslí opakovač
- má několik portů (aspoň dva) , kam se připojují jednotlivé sítě
- pracuje na úrovni bitů, neví, co jednotlivé bity znamenají
- bity jsou přijímány na jednom portu, vyhlazeny a opět rozesílány na všechny ostatní porty
 - opakuje data → opakovač
- dochází k opravě zkreslení a útlumu fyzických přenosových cest
 - regenerace signálu

Opakovač

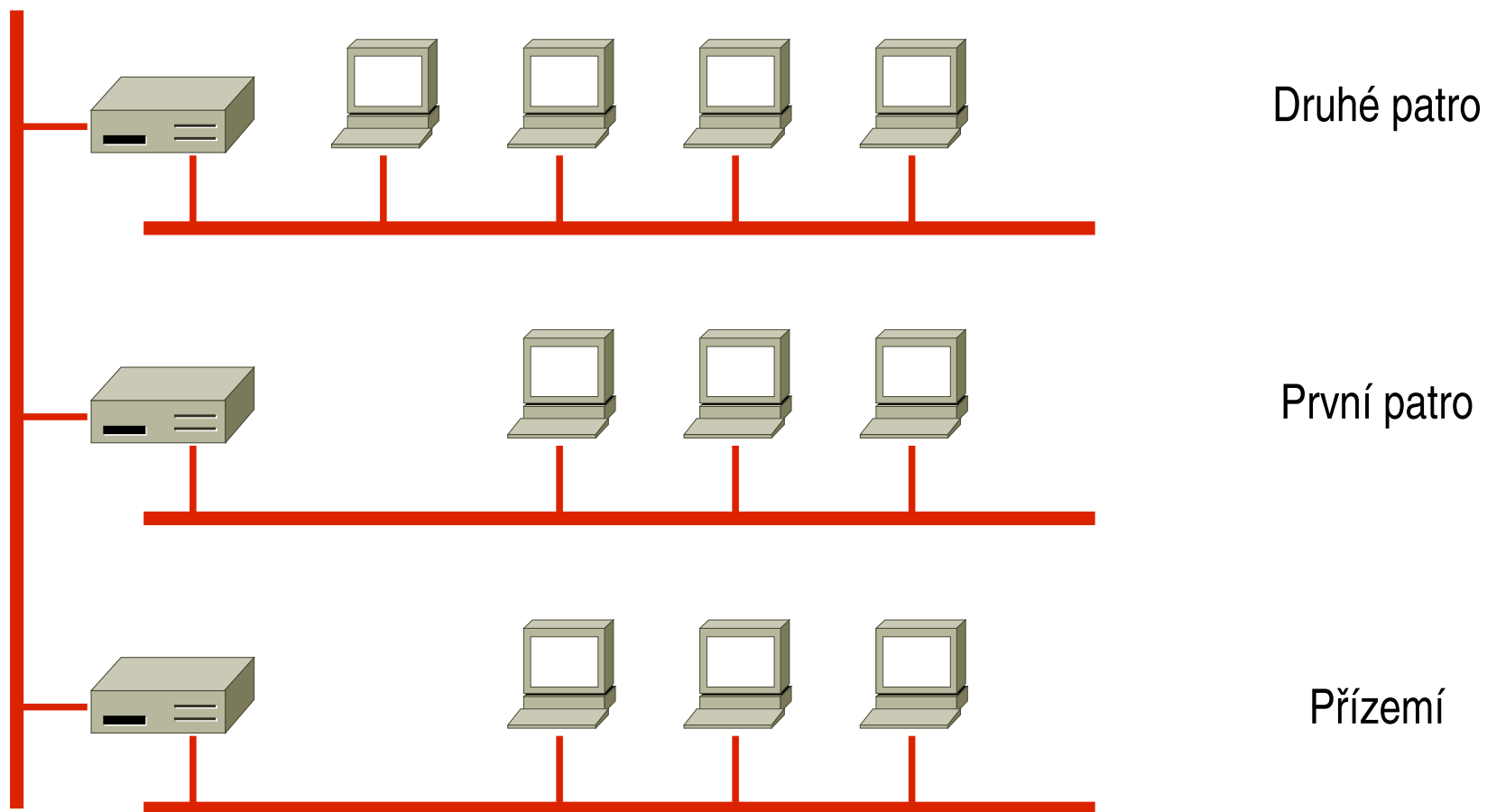
- opakovač způsobuje pouze malé přenosové zpoždění
 - max. několik bitů
- nemá žádnou vyrovnávací paměť
- může propojovat libovolný počet segmentů
- může propojovat pouze segmenty se stejnou rychlostí
 - bufferu nemaje



Opakovače v Ethernetu

- nemůže jich být v síti libovolně mnoho
 - vyžaduje metoda CSMA/CD – musí být zaručena určitá doba šíření, kvůli detekci kolizí
 - opakovač v Ethernetu musí šířit kolize(!)
 - všechny segmenty propojené pomocí opakovačů tvoří *kolizní doménu*
 - pravidlo: mezi každými dvěma body maximálně 2 opakovače
 - umožňuje páteřní rozvody
 - přesnější pravidlo: 5-4-3
 - 5 segmentů, 4 opakovače a 3 aktivní segmenty

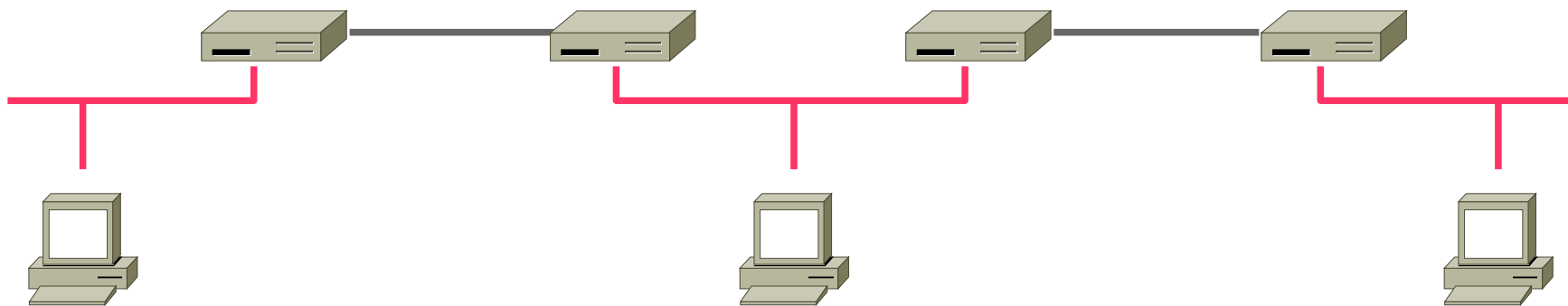
Opakovače v Ethernetu



Pravidlo max. 2 opakovače mezi dvěma uzly

Opakovače v Ethernetu

Pravidlo 5-4-3



mezi každými dvěma uzly sítě platí:

- max. 5 segmentů 4 opakovače a 3 aktivní segmenty

Filtrování

- opakovače jsou hloupé – šíří do všech směrů i to, co by nemusely
- chtěli bychom, aby propojovací zařízení poznalo, co musí přeposlat a co nikoliv
 - neboli filtrovat provoz
 - musí rozumět aspoň části přenášených dat – adresám linkové vrstvy
 - neboli musí pracovat na linkové vrstvě
 - pokud se bude rozhodovat podle adres, nemůže fungovat v reálném čase, musí si bufferovat rámce (nebo jejich části), aby bylo možné rozhodnout, na které porty mají být vyslány
 - díky bufferování ale může propojovat různě rychlé segmenty sítě

Most, přepínač

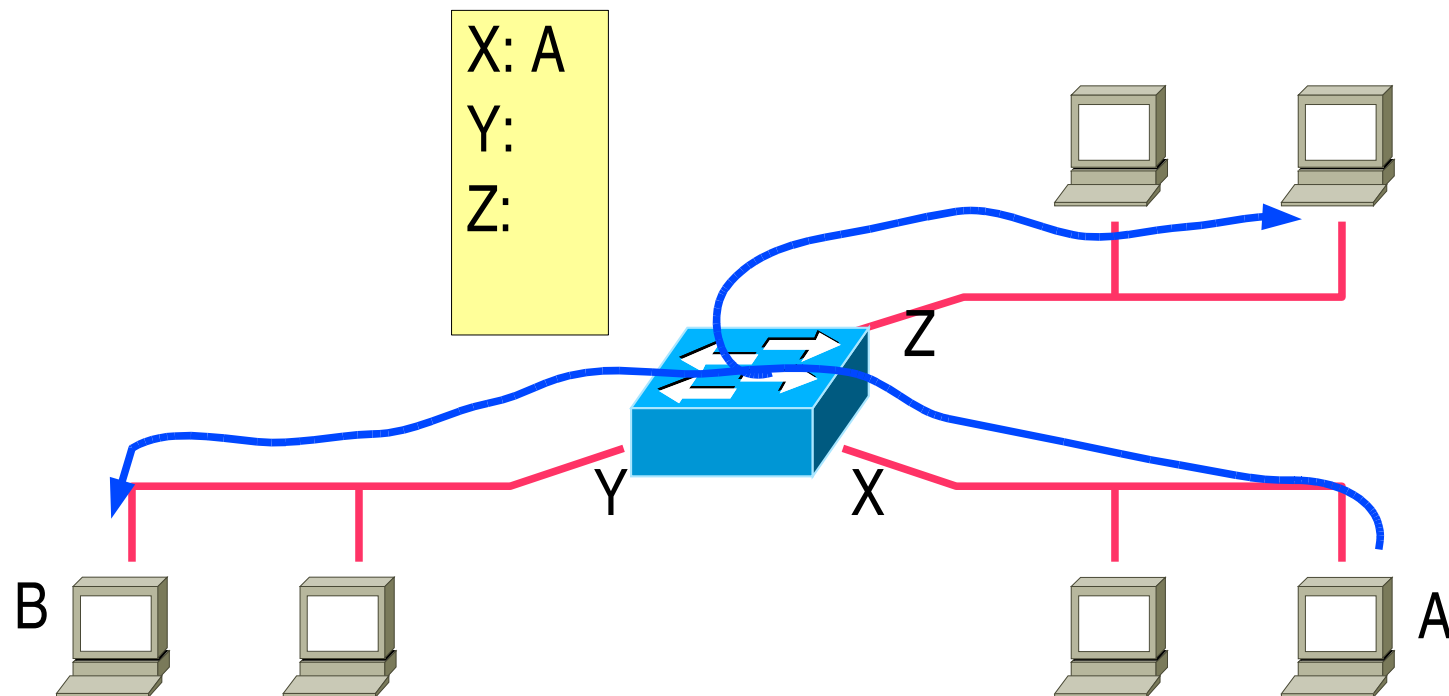
- díky bufferování nemusí propagovat kolize
 - odděluje kolizní domény!
- všesměrově (v ideálním případě) šíří pouze broadcasty
 - pakety určené všem příjemcům v dané síti
- most (bridge): starší, většinou bufferuje celé rámce, pomalejší, má pouze dva porty – propojuje dvě sítě
- přepínač (switch): novější, víceportový (8, 16, 24, 48), rychlý
 - často se používá jako náhrada hubu
- přes jeden přepínač může (na různých portech) probíhat několik přenosů (až do maximální kapacity přepínače)

Přepínač: filtrování

- aby most mohl filtrovat, musí znát topologii sítě
 - nebo aspoň na kterém portu se nachází která stanice
 - pokud neví, musí se chovat jako opakovač (ale na linkové vrstvě)
- můžeme na každém přepínači nastavit ručně
 - mnoho práce, může se měnit (např. notebook v síti, ...)
- switch si potřebné údaje zjistí sám
 - zpočátku se chová jako opakovač a učí se – ukládá si do tabulek dvojice port-adresa
 - po určité době začne filtrovat provoz
 - plug & play

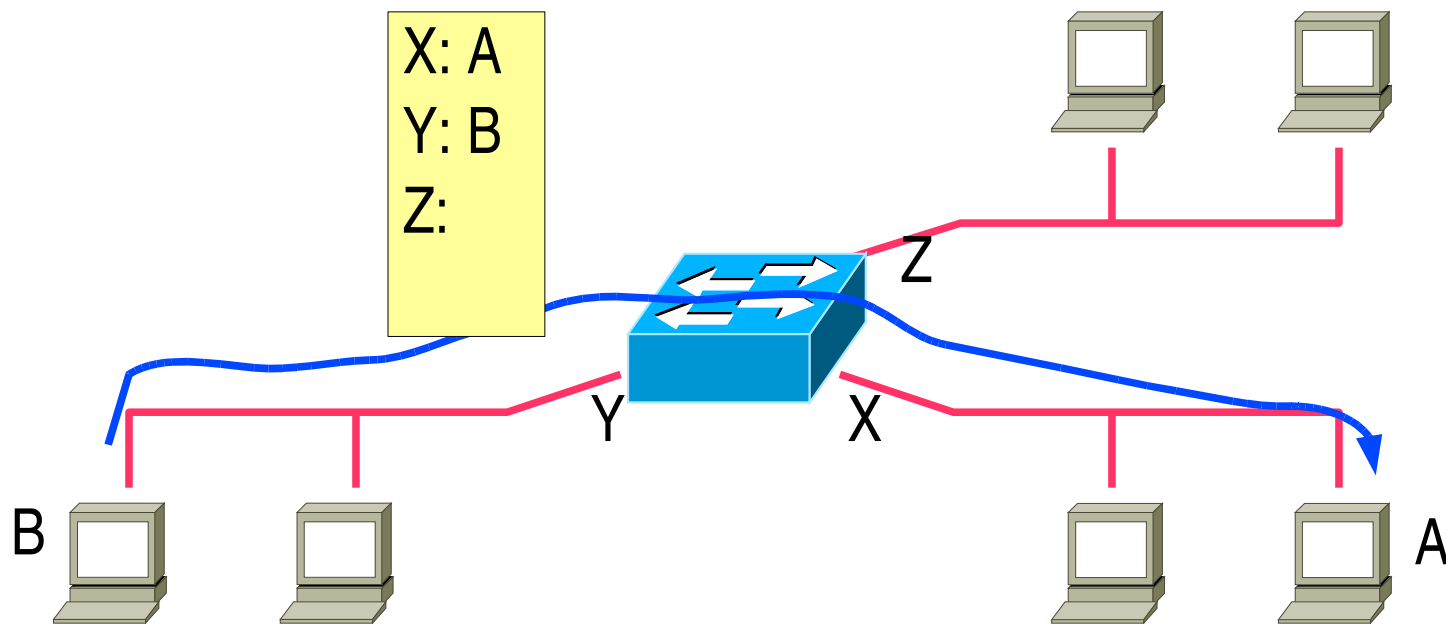
Učení

- když dostane rámeček „od A k B“ na portu X, odvodí, že uzel A je na portu X, zapíše si danou informaci do tabulky
- pošle rámeček na všechny porty kromě X



Učení

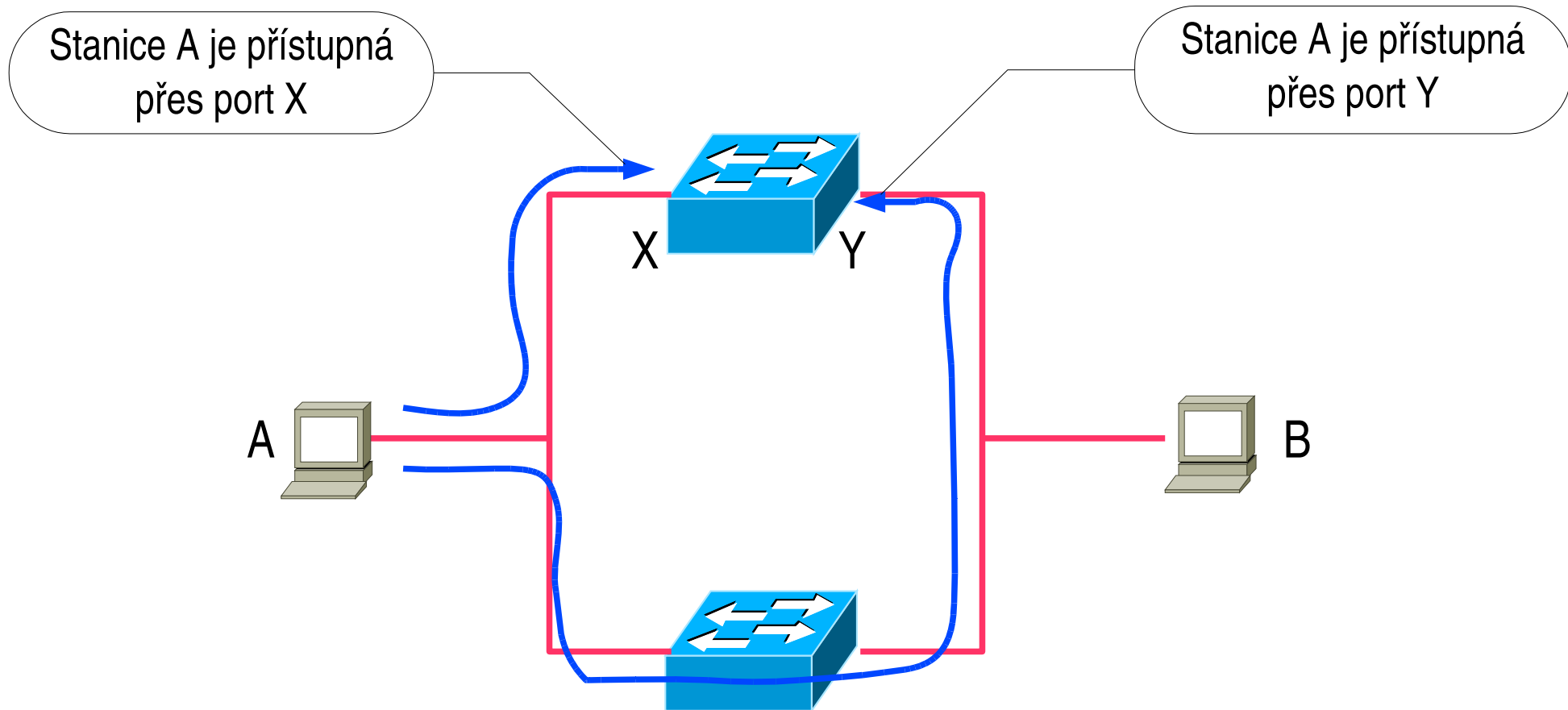
- stanice B odpoví, switch se z odpovědi dozví, kde leží B
 - kde leží stanice A už ví
- další rámce komunikace mezi A a B už posílá pouze mezi zjištěnými porty (X, Y)



Problém: kružnice

- učení přestane fungovat, pokud struktura sítě nebude stromová, ale bude obsahovat kružnici
- mosty se dokáží domluvit a kružnici přerušit
- kružnice může být v síti úmyslně: kvůli možnému výpadku spojení
- STA: Spanning Tree Algorithm
 - algoritmus, pomocí kterého se najde kostra dané síťové topologie
 - přeruší se případné kružnice (zablokují se některé porty)
 - zůstane stromovitá struktura
 - v případě výpadku se zablokované porty opět uvolní
 - všechny moderní switche algoritmus podporují (IEEE 802.1d)

Problém: kružnice



Spanning Tree Algoritmus

- nejprve se zvolí kořenový switch (root switch)
 - ten s nejmenší MAC adresou (případně se dá nastavit administrativně)
- na každém síťovém segmentu se zvolí designated switch
 - to je ten, který má nejlepší cestu ke kořenovému switchi
- porty switchů, které nejsou na kostře grafu sítě se přepnou do blokujícího stavu
- možné stavy portu: blocking, listening, learning, forwarding, disabled
 - blocking: neprocházejí rámce, slouží jako záložní
 - learning: neprocházejí rámce, switch se učí adresy
 - forwarding: procházejí rámce (a také se učí adresy), cílový stav pro aktivní porty v kostře grafu sítě

Source routing

- alternativní možnost: přepínání rámců stylem „source routing“
 - neboli zdrojové směrování, spíše technika síťové vrstvy
 - o tom, kudy bude rámeček sítě procházet rozhoduje odesílající uzel
 - do rámce zapíše adresy všech switchů, přes které rámeček projde
- jak odesílatel zjistí cestu?
 - pomocí záplavového rozesílání: pošle paket, který se rozešle všem sousedům
 - rekurzivně až k cíli, cíl odpoví a v odpovědi je cesta, kterou původní paket prošel
 - je to dosti nešetrné k síti
 - ale najde opravdu nejlepší cestu

Repeater, switch nebo router?

- neexistuje jednoznačná odpověď
 - ale existují doporučení
- chceme-li zvýšit propustnost sítě:
 - zvětšíme přenosovou rychlost (100 Mbps -> 1 Gbps)
 - např. jen u některých stanic/serverů
 - rozdělíme síť na více částí, s cílem maximalizace lokálního provozu
- udává se, že 80% provozu je lokálního charakteru, 20% mezi sítěmi
 - dneska neplatí, díky Internetu

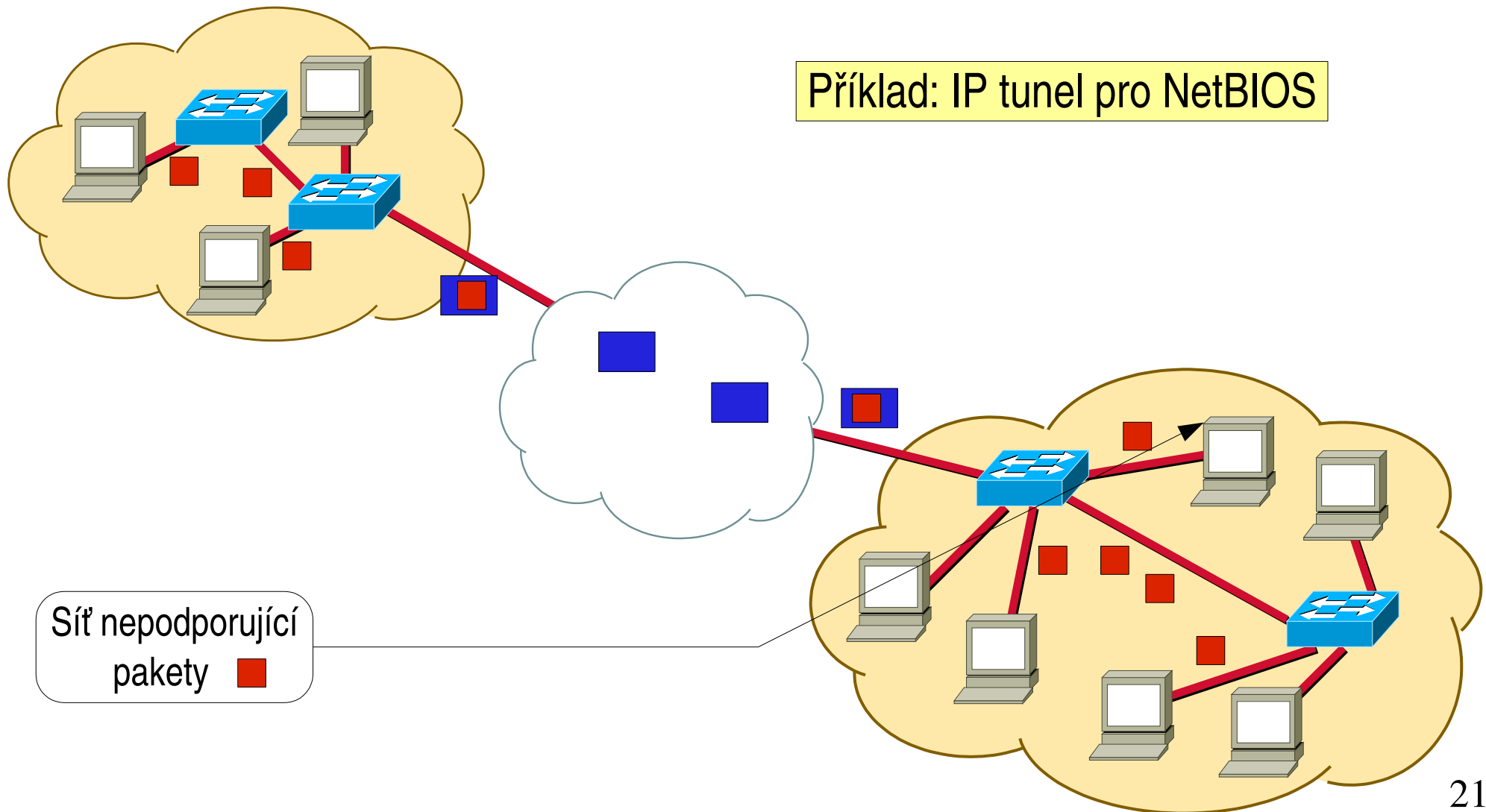
Repeater nebo switch?

- opakovač
 - neodděluje kolizní domény (šíří se skrz něj kolize)
 - je rychlý, ale počet v síti je omezen
 - nepodporují redundantní cesty
- přepínač
 - umožňuje vytvářet rozsáhlé sítě, je pomalejší
 - dnes se téměř všude používají switche (přepínače)
 - nepodporují redundantní cesty (ale umí se s nimi vyrovnat)
 - moderní switche lze nakonfigurovat pro agregaci – více linek se tváří jako jedna
 - ne vždy musí být nejvýhodnější

Jak spojit různé technologie

- opakovač (fyzická vrstva)
 - nelze: různé rychlosti, různá technologie
- přepínač (linková vrstva)
 - lze, pokud jsou sítě dostatečně podobné (např. TokenRing, Ethernet)
- router (síťová vrstva)
 - standardní řešení
- co s nesměrovatelnými protokoly? (NetBIOS)
 - neznají pojem sítě ... dají se použít jen v rámci jedné sítě
 - řešení: zapouzdření do jiného protokolu

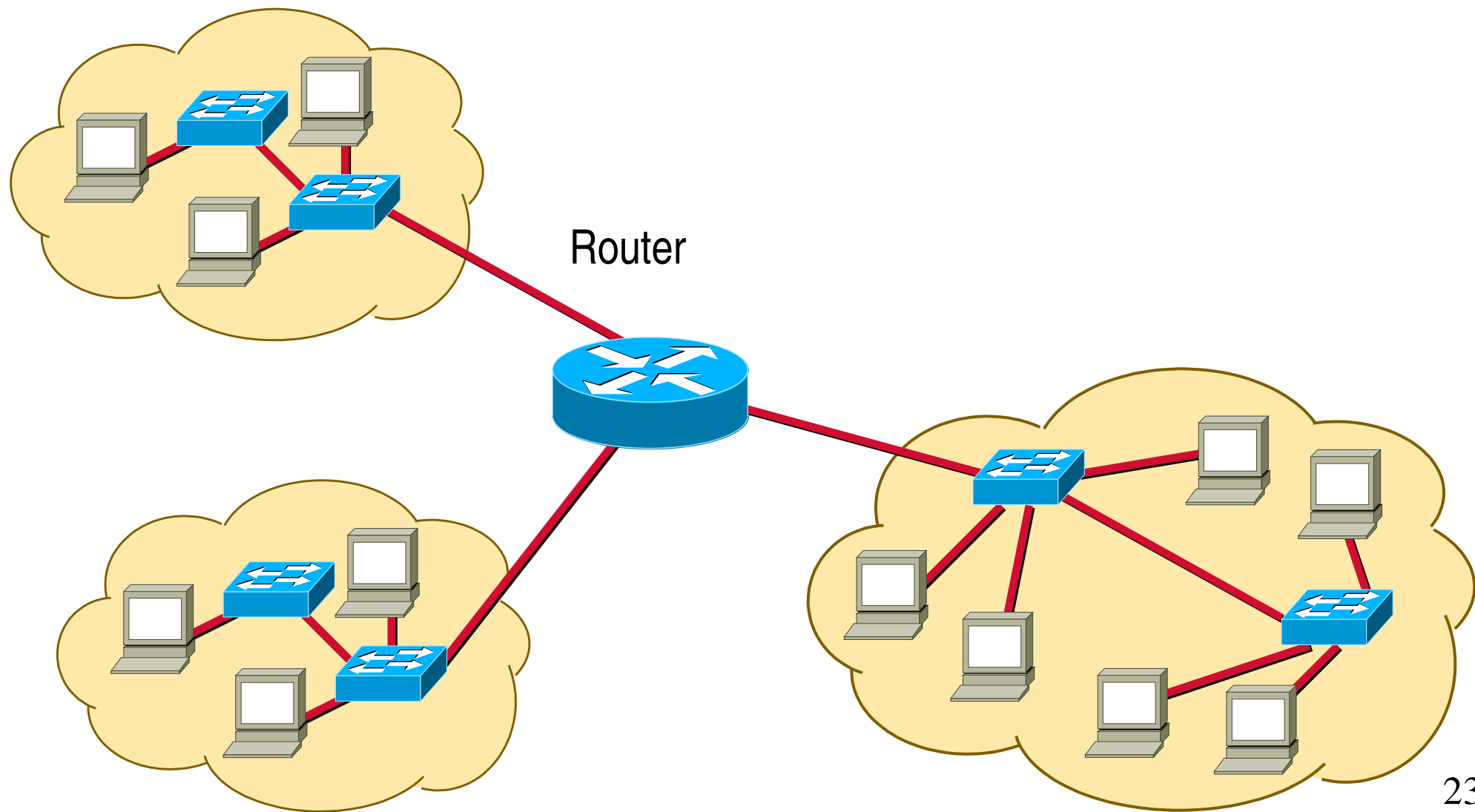
Tunelování (zapouzdření)



Switch nebo router?

- přepínač
 - vytváří jednu velkou síť, nemá explicitní požadavky na maximální velikost sítě
 - v rámci sítě mohou uzly komunikovat „přímo“, bez prostředníka
 - šíří broadcasty v celé síti – je vhodné ji omezit
- směrovač (router)
 - pracuje na síťové úrovni, propojuje jednotlivé sítě
 - nešíří broadcasty z jedné sítě do druhé
- kritérium pro rozdělení do sítí:
 - lokální:nelokální = 80:20
 - snadná správa, přístupová práva

Klasické řešení



Switche dnes

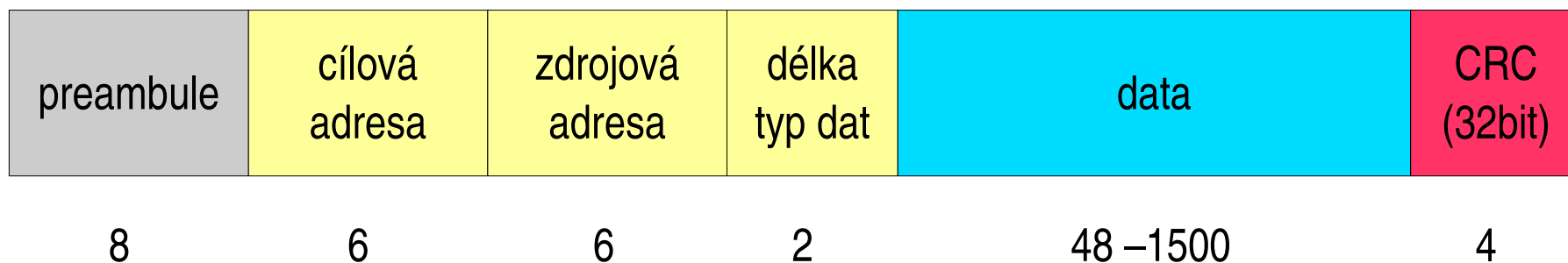
- podporují QoS (quality of service), VLAN (Virtual LAN)
- různé verze STA (s rychlejší konvergencí, ...)
- agregaci více linek do jedné
- filtrování rámců na různých úrovních:
 - linková: Layer 2
 - síťová: Layer 3, filtrování na úrovni protokolu IP, IP adresy...
 - transportní: Layer 4, filtrování na úrovni portů TCP/UDP
- Layer 3 switch: vlastně router, ale optimalizovaný na rychlost, HW řešení
- Layer 4 switch: umožňuje rozlišovat provoz, např. load balancer

VLAN

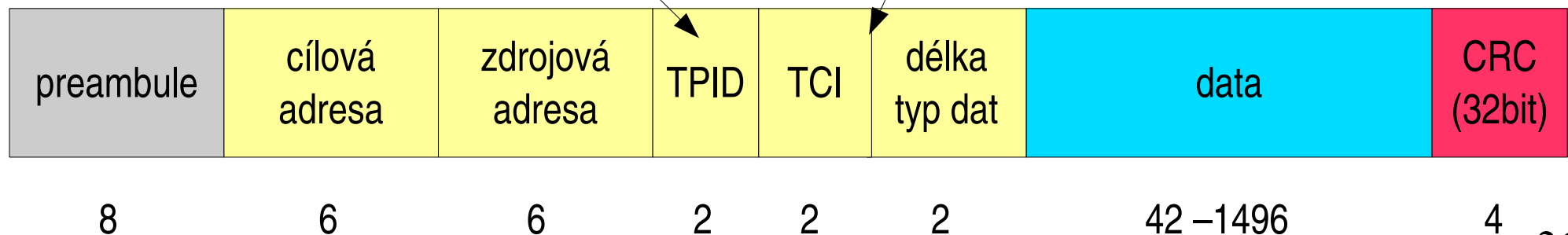
- Virtuální síť, standard IEEE 802.1q
- VLAN je skupina uzlů, které komunikují, jako by byly spojené LAN
 - ale jsou v jiných LAN nebo tvoří v LAN podmnožinu
 - IEEE 802.1q umožňuje vytvářet VLAN, které tvoří „podsít“ v rámci jedné LAN
- funguje na principu značkování rámců
 - pro každý port switche je nastaveno, do které VLAN patří
 - příchozí data se označují značkou příslušné VLAN a pošlou se dále
 - switch má různé tabulky pro virtuální síť
 - data se před odesláním ze switche cílovému uzlu odznačují (většina koncových zařízení nepodporuje IEEE 802.1q)

IEEE 802.1q: rámce

IEEE 802.3



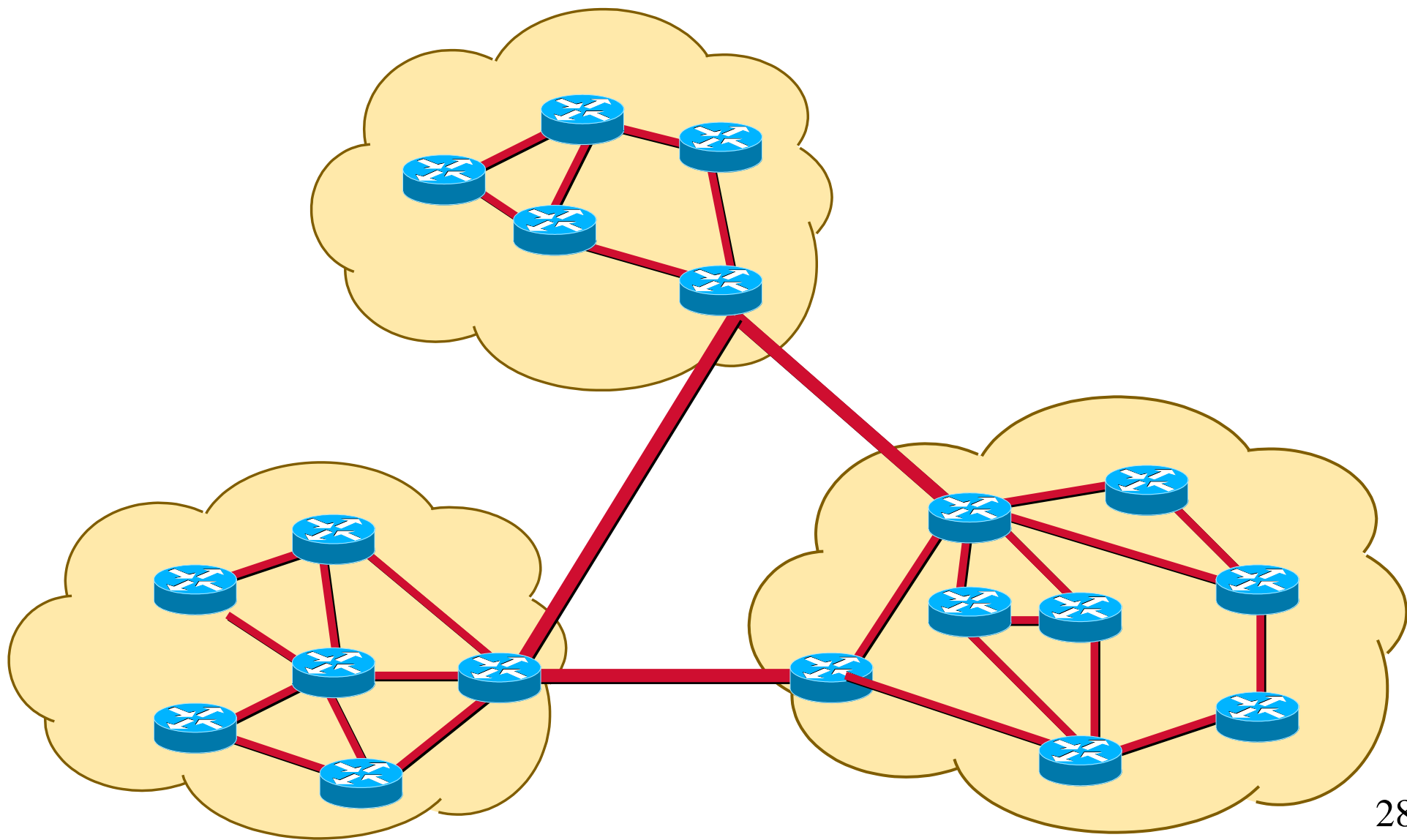
IEEE 802.1q



Autonomní systémy

- směrování můžeme provádět nejrůznějším způsobem
 - statické tabulky (vhodné pro malé systémy)
 - dynamicky: používáme speciální algoritmy pro distribuci cest v síti
- směrovací algoritmy se mohou lišit, mají mnoho parametrů
- v každé síti můžeme požadovat jiné vlastnosti směrování
 - kdo kam smí, kudy se bude směřovat provoz
 - definujeme „směrovací politiku“
- Autonomní systém (AS): několik sítí se společnou politikou směrování
 - většinou síť jednoho vlastníka (typicky ISP)

Autonomní systémy



Autonomní systémy

- AS používají jednotnou strategii směrování
 - v Internetu IGP: Interior Gateway Protocols
 - OSPF, RIP, IGRP
 - jednotná strategie pro směrování do ostatních AS
- AS jsou vzájemně propojeny
 - EGP: Exterior Gateway Protocols – protokoly pro směrování mezi AS
 - v Internetu se používá BGP (Boarder Gateway Protokol)
- peering: propojení několika AS
 - data pak neprocházejí přes jádro Internetu, ale lokálně
 - v ČR: NIX.CZ, propojuje desítky sítí