



RNDr. Ing. Vladimír Smotlacha, Ph.D.

Katedra počítačových systémů
Fakulta informačních technologií
České vysoké učení technické v Praze
© Vladimír Smotlacha, 2019

Počítačové sítě BI-PSI

LS 2018/19, Přednáška 5

<https://courses.fit.cvut.cz/BI-PSI>



FAKULTA
INFORMAČNÍCH
TECHNOLOGIÍ
ČVUT V PRAZE



EVROPSKÁ UNIE

Směrování

Propojování sítí

- opakovač (hub, repeater)
- most (bridge)
- přepínač (switch)
- směrovač (router)
- brána (gateway)

Autonomní systémy

Směrovací protokoly

- RIP
- OSPF
- BGP



- záplavové
- náhodné
- statické
- dynamické



Směrovač odešle přijatý paket na každou výstupní linku

- doručení v nejkratším možném čase
- omezená životnost paketů – čítač v hlavičce
- paket se duplikuje exponenciálně
 - vylepšení: směrovač si pamatuje paket a zpracuje ho jen jednou
- velmi neefektivní využití sítě



Přijatý paket se odešle do náhodně zvolené linky

- nezaručuje konečnou dobu doručení
 - v základní podobě není reálně použitelný
- lze využít jako doplněk k jiným algoritmům
 - při zahlcené výstupní lince je paket odeslán do jiné náhodně zvolené linky
 - reálná šance, že bude doručen



Směrovací tabulka je fixní

- vytvořena podle konfigurace sítě
- nemění se podle stavu sítě
- neschopnost reagovat na poruchy sítě
- příklad: počítač v lokální síti
 - dva záznamy v tabulce
 - adresy v lokálním segmentu sítě
 - default gateway – pro všechny ostatní adresy



Směrovací tabulka se mění podle stavu sítě

- způsob aktualizace
 - izolovaně
 - směrovač provádí změny samostatně bez ohledu na ostatní směrovače
 - centralizovaně
 - výpočet nových tabulek prováděn centrálně
 - tabulky jsou distribuovány do směrovačů
 - decentralizovaně
 - výpočet provádí každý směrovač
 - pokladem jsou data od ostatních směrovačů

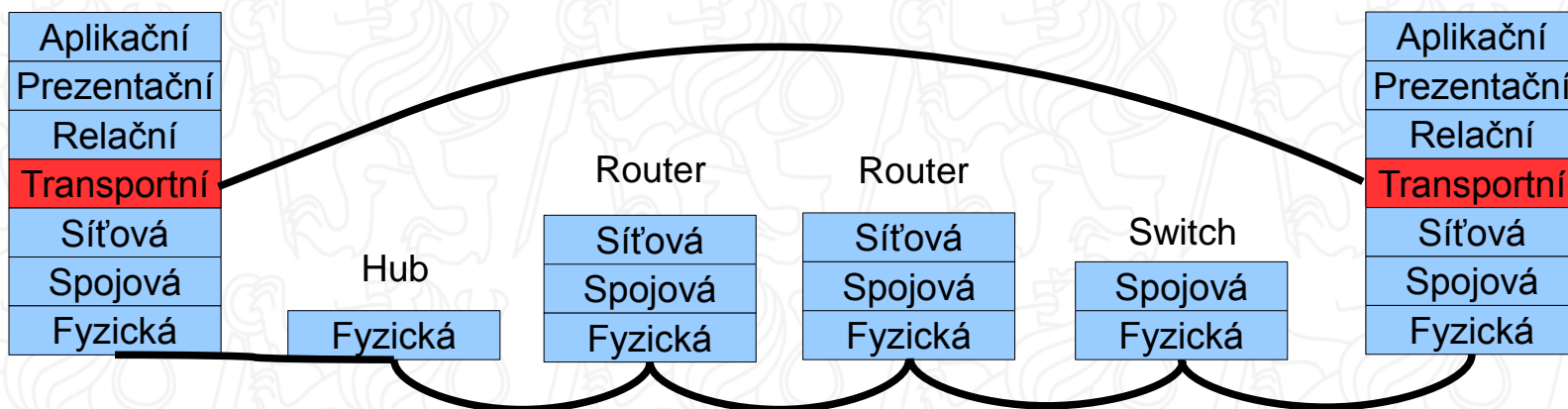


- DVA – Distance Vector Algorithm
 - směrovače periodicky vysílají obsah svých tabulek sousedům
 - aktualizace vlastní tabulky, pokud je nalezena „kratší“ cesta
 - metrikou je počet uzlů na cestě
 - problém: omezený „průměr“ sítě
 - vše ve větší vzdálenosti je „nekonečno“ a tedy nedosažitelné
- LSA – Link State Algorithm
 - směrovače se vzájemně informují o stavu linek
 - každý směrovač má úplnou informaci o topologii sítě
 - pomocí Dijkstrova algoritmu spočítá mapu nejkratších cest



Komunikace se vztahuje se k určité vrstvě n

- v síťovém prvku implementován stack až do n
- funkce
 - přijatý blok dat (paket, rámeček) je předán do vrstvy n
 - podle obsahu hlavičky se určí akce
 - datový paket putuje stackem zpět a je odeslán





Pracuje na fyzické vrstvě

- neobsahuje žádnou paměť ani složitou logiku
- repeater: 2 porty , hub: více než 2 porty
- zesiluje signál
 - zvětšení vzdálenosti mezi komunikujícími uzly
 - překonání útlumu
 - zlepšení odstupu signálu od šumu
- v kolizní síti (např. CSMA/CD) musí šířit kolize
- nelze překročit danou velikost segmentu



Pracuje na linkové vrstvě (L2)

- neplatí omezení na fyzickou velikost segmentu
- obsahuje vyrovnávací paměť
- bridge: 2 porty, switch: více než 2 porty
- propojuje dvě sítě, přenáší rámce z jedné sítě do druhé
 - využívá adresy MAC
- odděluje kolizní segmenty
- moderní, nejčastěji používaný aktivní síťový prvek
 - v LAN nahradil rozbočovače (hub)
 - široký rozsah typů: 10Mb/s – 10Gb/s, 4 – 100+ portů



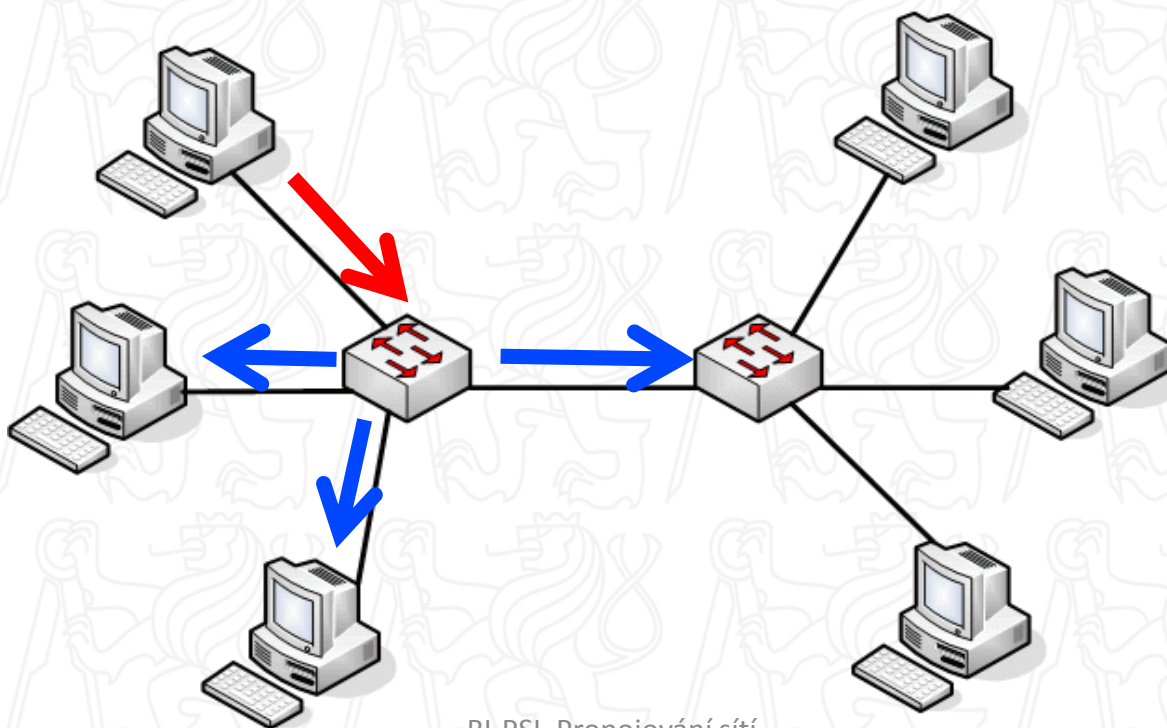
- Pamatuje si přiřazení adres MAC k portům
 - tabulka dvojic [port, adresa]
 - sám se učí podle MAC odesílatele (source MAC)
 - záznamy po určité době expirují
- funkce
 - rámec se známou cílovou MAC odešle jen na daný port
 - rámce s dosud neznámou MAC odešle na všechny porty
 - broadcast MAC (FF:FF:FF:FF:FF:FF) odešle na všechny porty
- snížení zátěže linek
- vyšší bezpečnost – data nelze odposlouchávat na každém portu jako u hubu



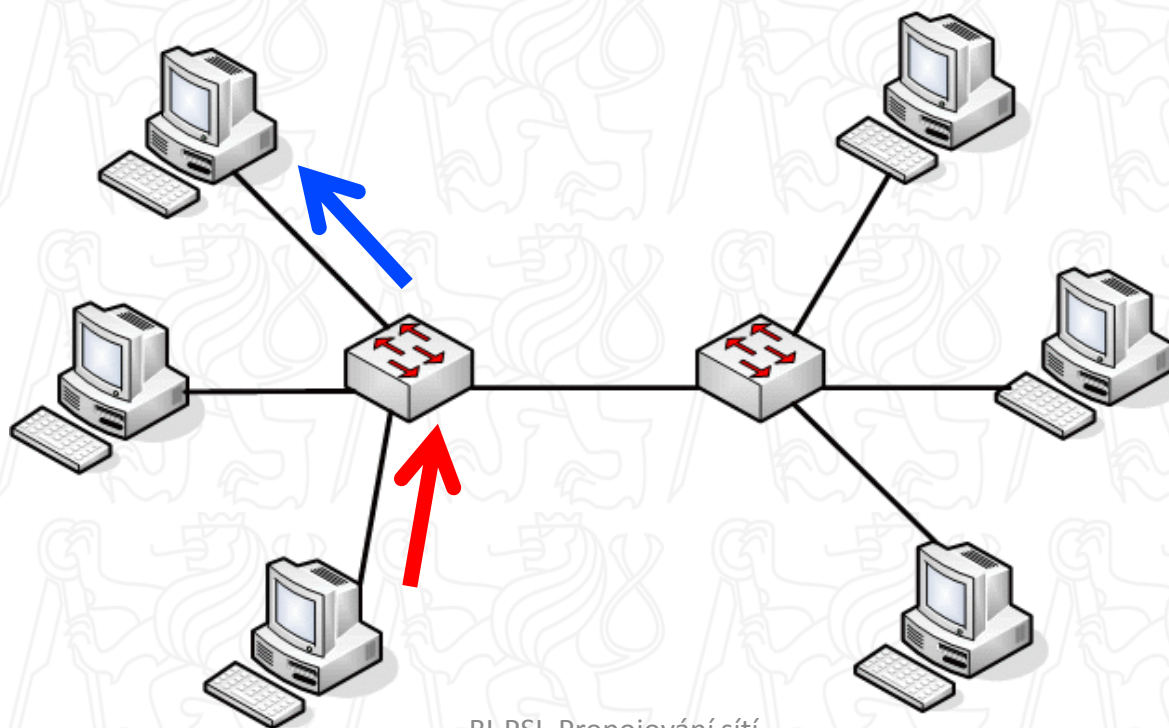
Dvě způsoby činnosti

- store-and-forward
 - celý rámec je napřed přijat a až poté analyzován a odeslán
 - zvyšuje se zpoždění pro delší rámce
 - poškozené rámce (např. kolize) se mohou hned zahodit
- cut-through
 - rámec je průběžně analyzován, je odeslán hned po přijetí cílové adresy
 - nižší zpoždění
 - i poškozené rámce se přenášejí
- v obou případech je potřebná paměť na uložení většího počtu rámců

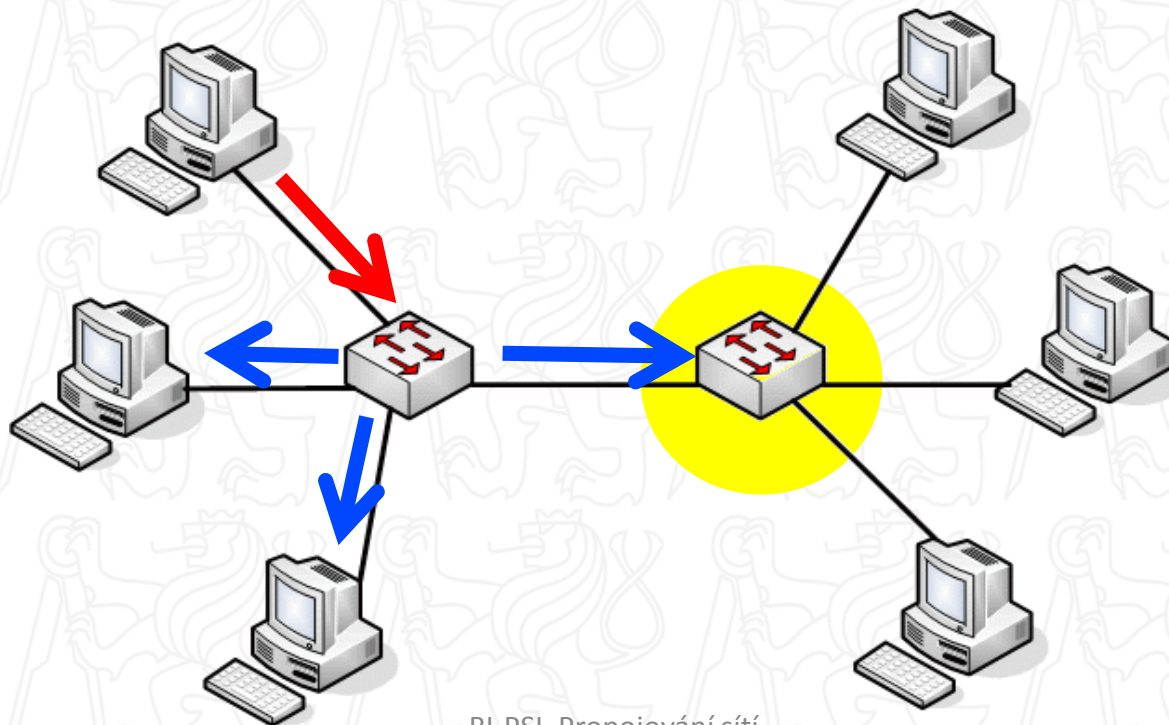
- přepínač dostane rámeček, o jehož cílové adrese nemá záznam
- pošle jej tedy do všech portů kromě toho, z kterého přišel, a upraví tabulku – už ví, kde je odesílatel



- přepínač dostane rámeček, o jehož cílové adrese již má záznam
- pošle jej tedy do portu uvedeného v tabulce pro danou adresu



- přepínač přijme rámec, který přichází ze směru, ve kterém podle jeho tabulky leží adresát
 - přepínač tento rámec ignoruje





- topologie switchů musí tvořit strom
 - z každého uzlu musí být jen jedna cesta ke jinému uzlu
 - budou-li dvě cesty, algoritmus selže
 - v tom případě totiž obsahuje graf kružnici
- problém: jak kružnici logicky „rozpojit“
 - Spanning tree algorithm
- záměrně vytvořené „kružnice“ na fyzické úrovni
 - nadbytečné linky mezi switchi
 - „horká záloha“ pro případ výpadku linky



- nalezne kostru dané topologie
- přeruší kružnice zablokováním některých portů
 - obnoví stromovou strukturu
- v případě výpadku některé linky aktivuje zablokované porty
 - algoritmus podporují všechny současné přepínače
 - specifikováno v IEEE 802.1d



- nejdříve se zvolí *kořenový přepínač* (root switch)
 - ten který má nejmenší MAC adresou
 - každý přepínač tvrdí o sobě, že je root, pokud neví o jiném kořenovém přepínači
- rozesílají HELLO MESSAGE
(Bridge Protocol Data Unit – BPDU)
 - BPDU jsou přeposílána přepínači se zvýšenou hodnotou cost (cena cesty ke kořenovému přepínači)

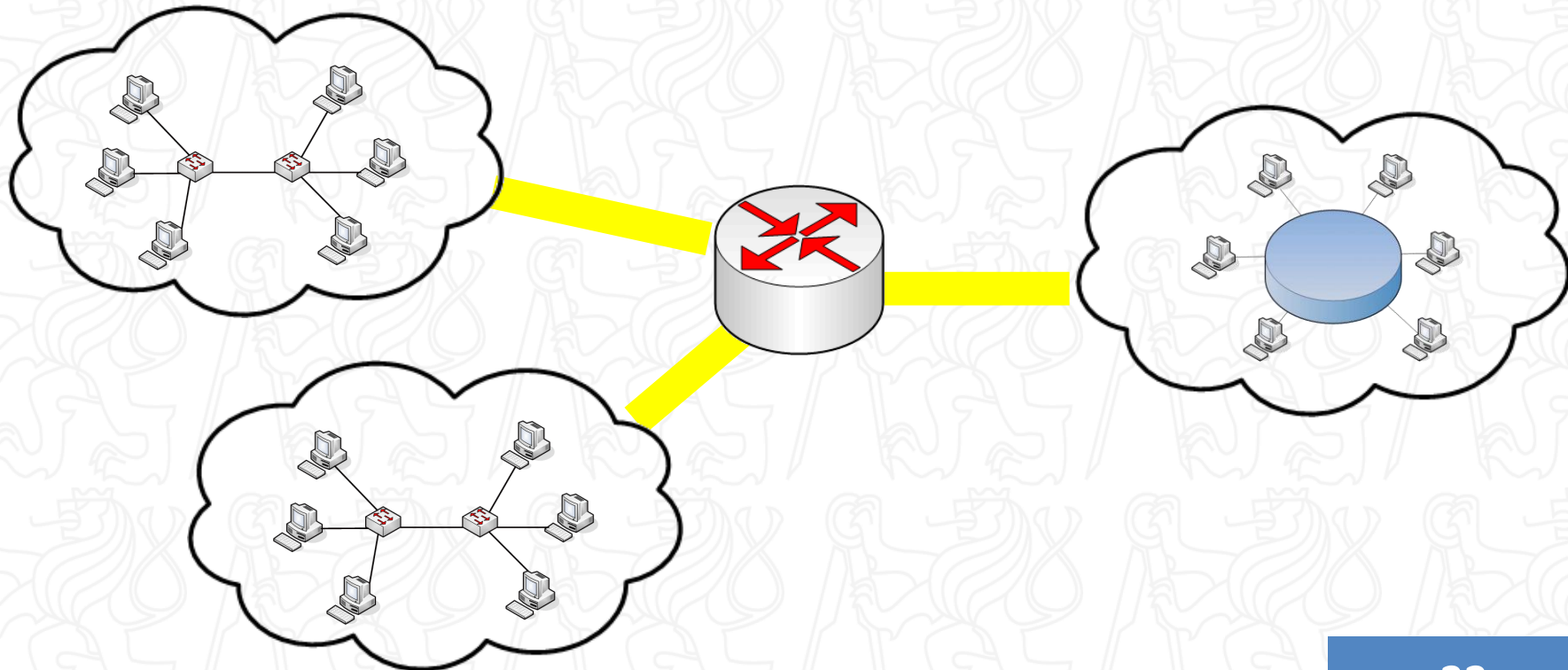


- na každém segmentu se zvolí vyhrazený přepínač (designated switch)
 - má nejlepší cestu ke kořenovému přepínači
- porty přepínačů, které nejsou na kostře grafu, se přepnou do blokujícího stavu
- konvergence STP při změně topologie trvá implicitně 50 sekund



- přepínače pracují s rámci, ale prohlížejí i hlavičky vyšších vrstev
- podporují QoS, VLAN, a další funkce
- zpracování rámců na různých úrovních:
 - linková
 - síťová => Layer 3 Switch (filtrování IP adres, IP protokolů)
 - transportní => Layer 4 Switch (filtrování na úrovni portů TCP a UDP, load balancer, ...)
 - aplikační => Layer 7 Switch (rozložení zátěže mezi více serverů)

- pracuje na síťové vrstvě (L3), propojuje sítě
 - využívá síťové adresy
- nešíří broadcasty z jedné sítě do druhé





- přenáší data i mezi sítěmi, které používají naprosto odlišné linkové technologie
 - jsou tedy naprosto nezbytné pro Internet, který má heterogenní strukturu
- kromě směrování může obsahovat i další funkce (firewall, NAT, VPN)
- směrování je statické nebo dynamické (např. RIP, OSPF, ...)



- Pracuje na vyšších vrstvách
- *Přenosová brána* je v rámci sítí obecný termín:
 - aplikační brána
 - směrovač bývá někdy také nesprávně označen jako brána
 - brána pro překlad protokolů z jedné množiny protokolů do jiné
- příklady aplikační brány:
 - software pro poštovní aplikace
 - proxy WWW server



Typický „WiFi router“

- univerzální kompaktní zařízení pro domácí síť nebo malou kancelář
 - 1 – 2x WiFi přístupový bod (access point) 802.11a/b/g/n
 - L2 switch (typicky 4 porty LAN, 100BASE-Tx nebo 1000BASE-Tx)
 - port WAN (100BASE-Tx)
 - L3 router (mezi porty WAN, LAN a WiFi)
 - DHCP server
 - podpora NAT
 - schopnost filtrovat provoz na L3 a L4



AS – autonomní systém

- IP sítě (skupiny sítí) pod společnou správou
- je identifikován číslem ASN – Autonomous System Number
- 16-bitů (původní schéma), 32-bitů (nové značení)
 - příklad: AS2852 – síť Cesnet2, ČVUT má privátní AS
- veřejný AS
 - ASN je celosvětově unikátní
- privátní AS
 - vyhrazen rozsah 64512 – 65535
 - privátní AS se nepropaguje do Internetu
 - číslo je viditelné jen v rámci sítě providera
 - navenek je nahrazeno číslem AS providera



2-úrovňová hierarchie

- směrování uvnitř autonomních systémů
 - IGP – Interior Gateway Protocols
 - např. OSPF, RIP
- směrování mezi autonomními systémy
 - EGP – Exterior Gateway Protocols
 - např. BGP



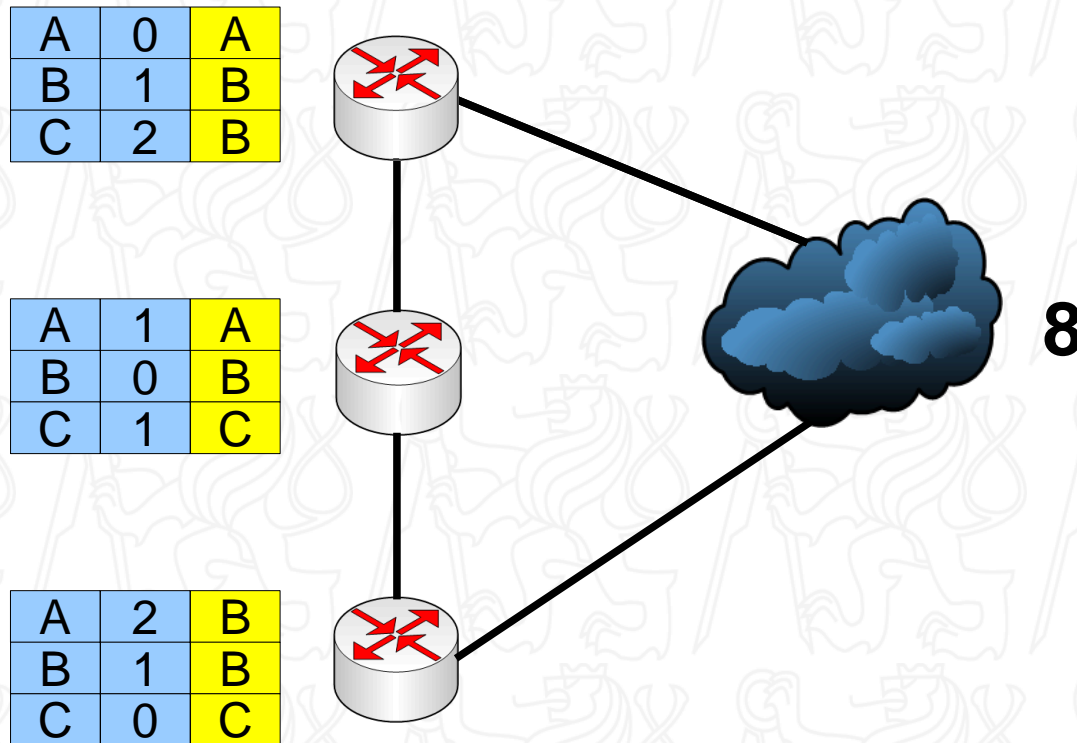
- Routing Information Protocol
 - RIP (RFC1058) – classful (class A,B,C)
 - RIPv2 (RFC1388) – classless (CIDR)
 - RIPv6 (RFC2086) – přidáno IPv6
- implementace DVA – Distance Vector Algorithm
 - Bellman-Fordův algoritmus hledání nejkratší cesty
 - metrikou je počet „skoků“ k cílové síti (hop count)
 - max. počet skoků (průměr sítě, tedy délka nejdelší cesty) je 15
 - 16 je už „nekonečno“ :-)

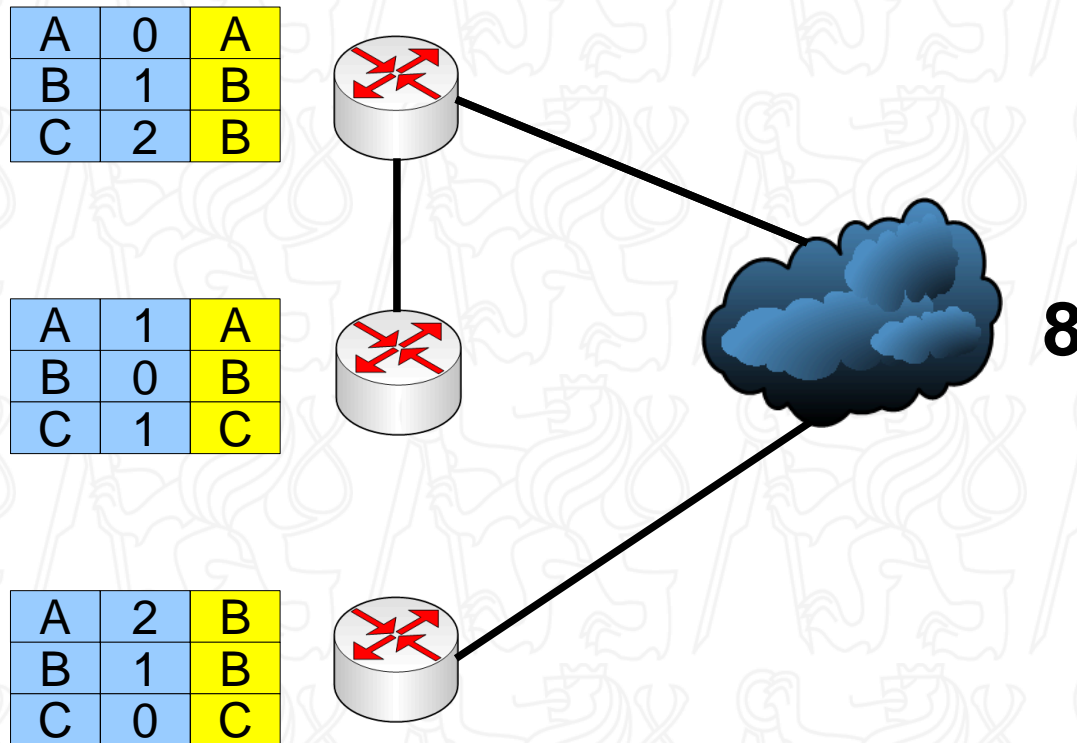


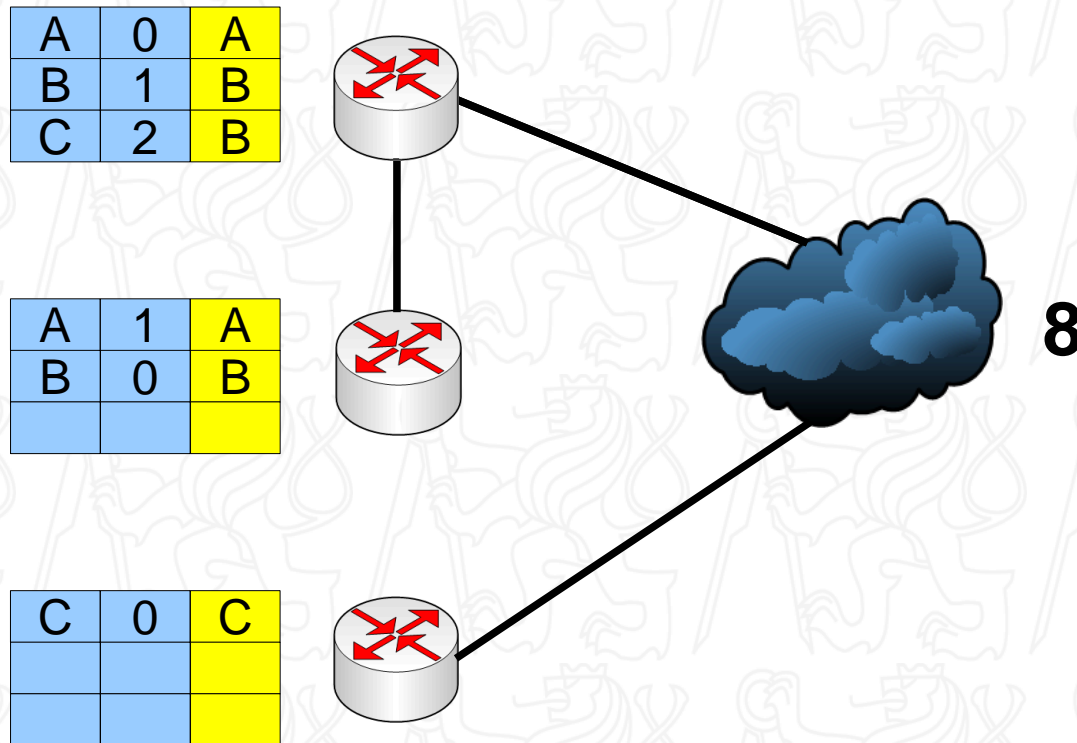
- tabulka trojic {Net ,G, D}
 - cílová síť (NET - Destination Network)
 - následující router (G - gateway)
 - vzdálenost (D - distance)
- inicializace: záznamy pro přímo připojené sítě
 - D=0, G = vlastní id
- router periodicky (každých 30 s) posílá svoji tabulku sousedům

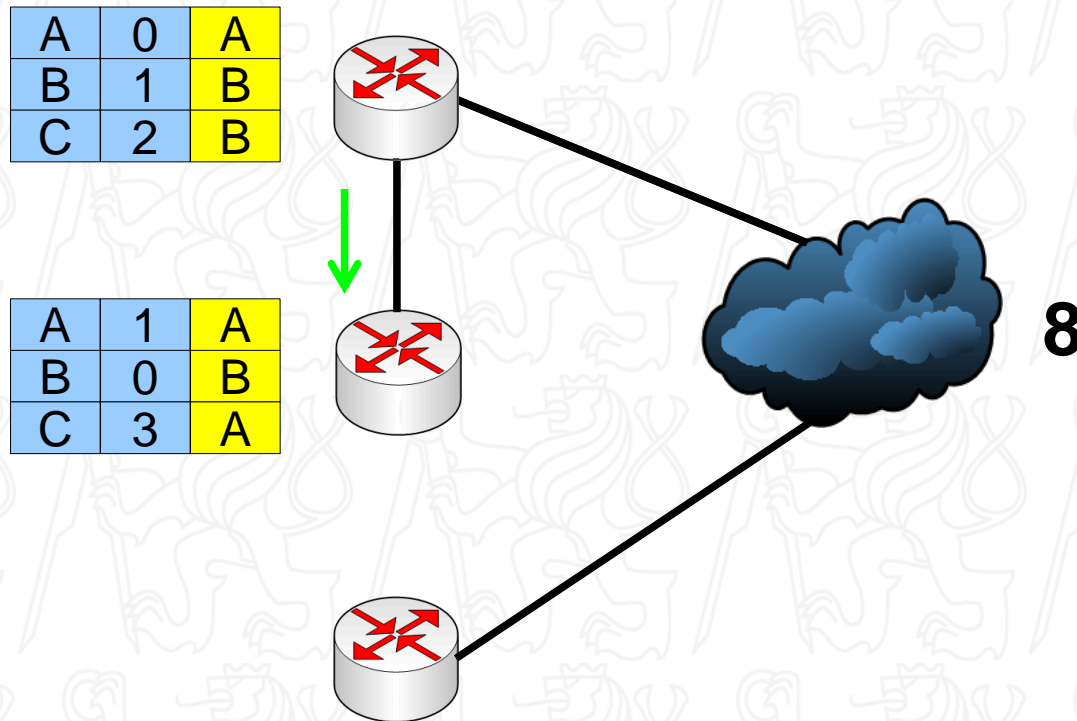


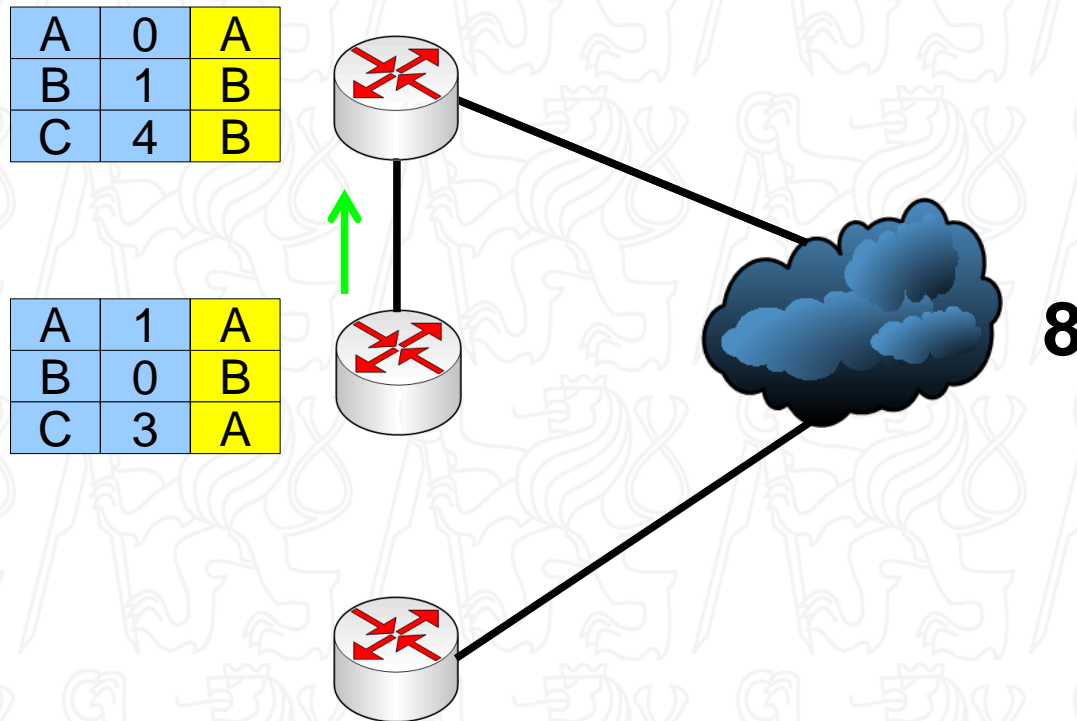
- obdržím tabulku souseda X {X:Net, X:G, X:D}
- aktualizace vlastní tabulky:
 - *Net* dosud neznámá
 - zapíše se {Net, X, X:D+1}
 - *Net* již existuje, $D > X:D+1$
 - nahradí se záznamem {Net, X, X:D+1}
- problémy RIP
 - omezená velikost sítě
 - zátěž sítě při předávání tabulek
 - pomalá konvergence po výpadku linky

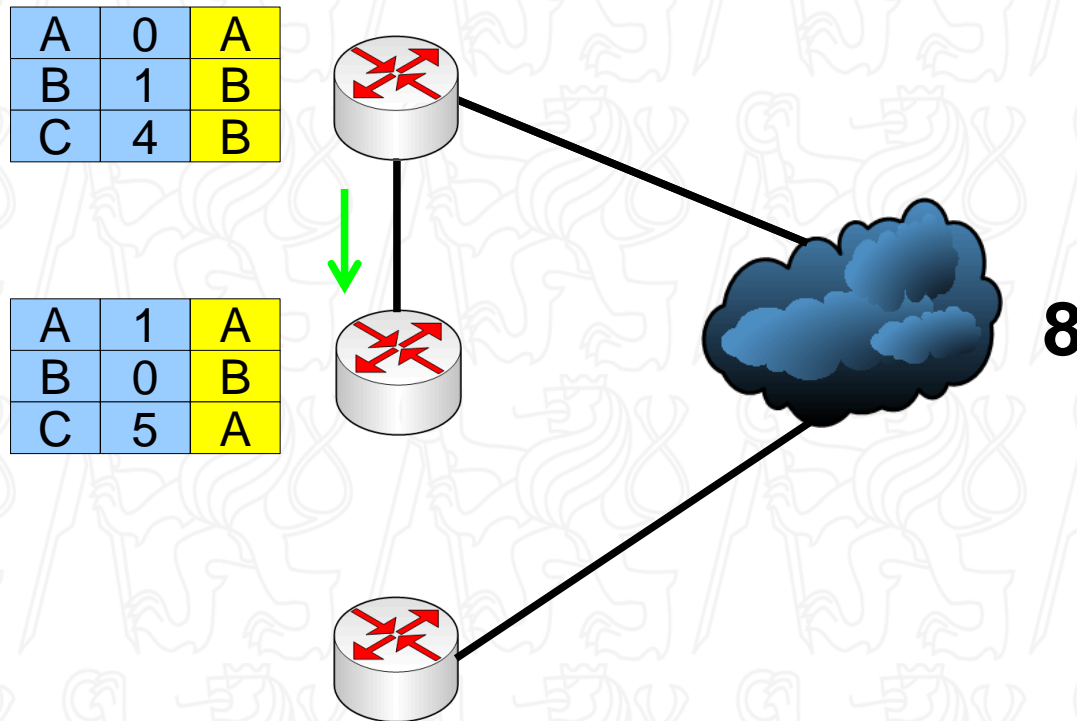


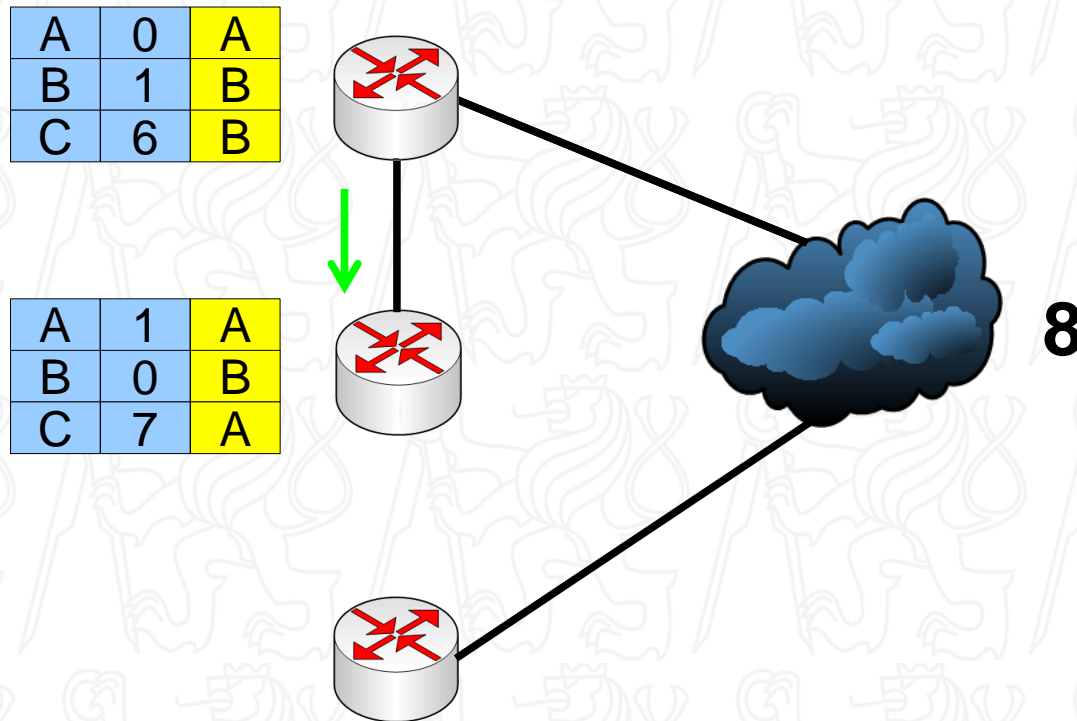


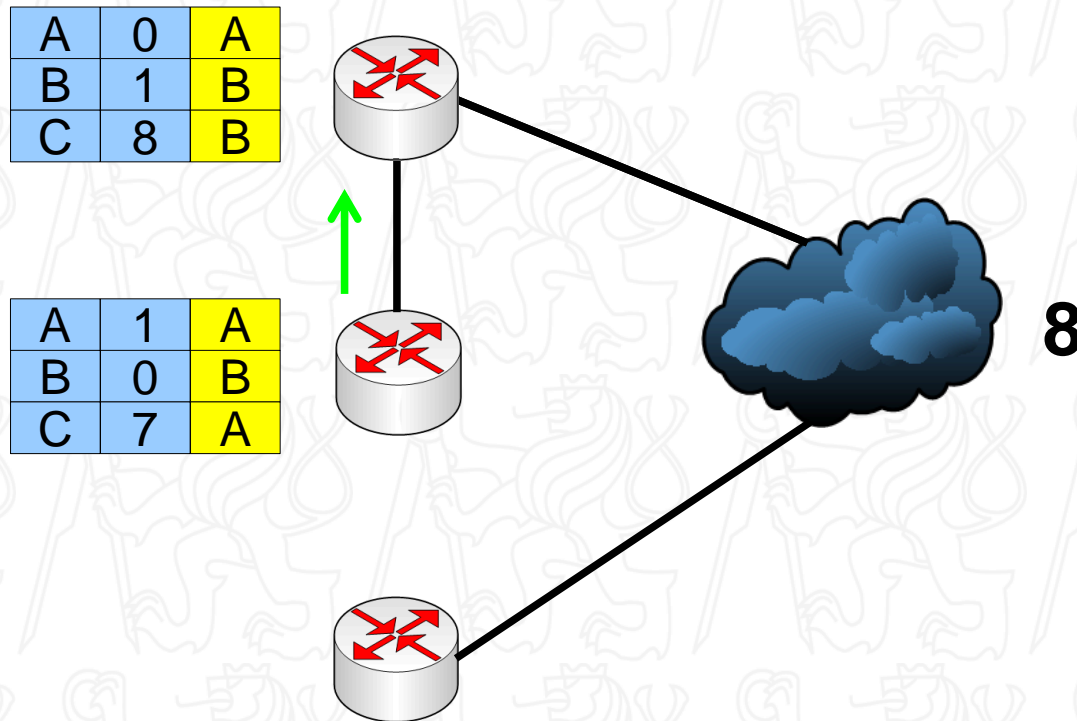


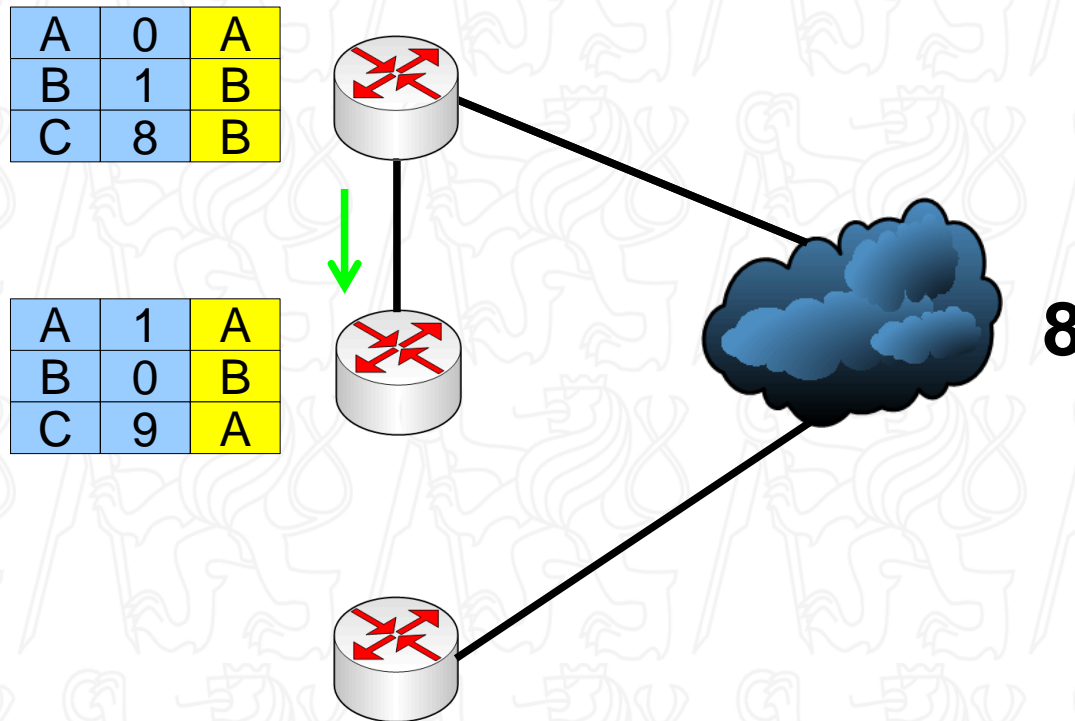


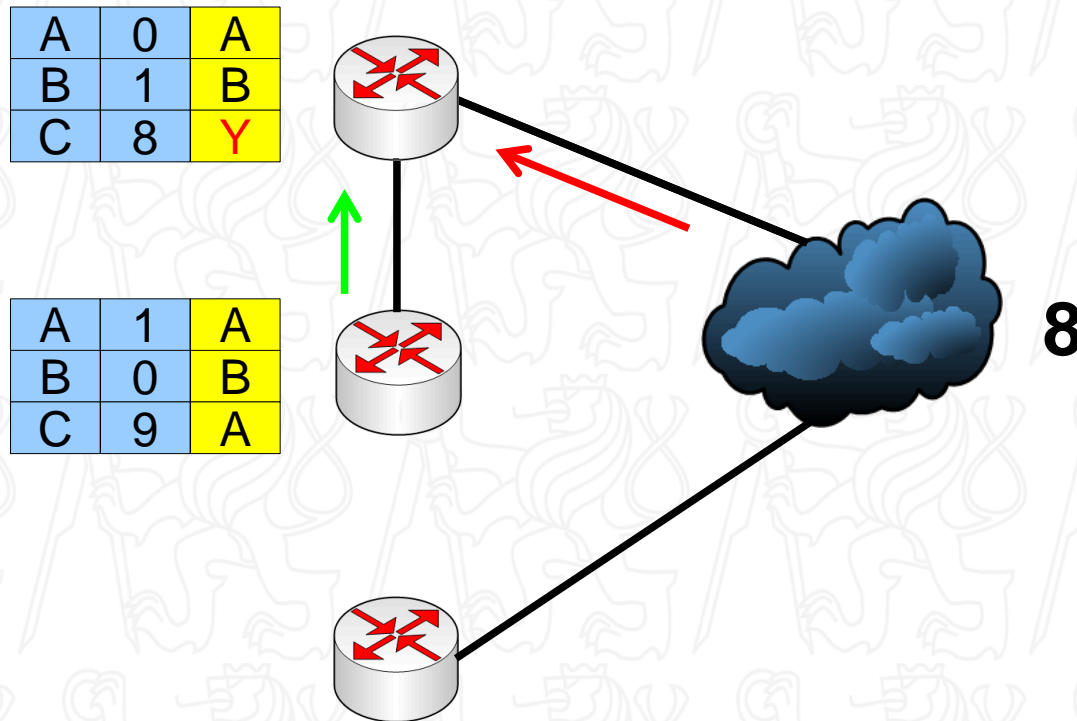














- Open Shortest Path First
 - OSPF v2 (RFC2328) – IPv4
 - OSPF v3 (RFC5340) – IPv6
- dominantní interní routovací algoritmus (Interior Gateway Protocol - IGP)
- implementace LSA – Link State Algorithm
 - Dijkstrův algoritmus nejkratší cesty v grafu
 - routery rozesílají stav linek (při změně, jinak 30 minut)
 - metrikou je skutečná kapacita linky
 - rozlišuje se typ služby
 - více možných cest v závislosti na službě



- každý router zná topologii sítě
 - přiřazení uzlů a hran
- každý router nezávisle spočítá tabulku nejmenších vzdáleností do všech uzlů
- rychlá konvergence (jednotky sekund)
- není striktní omezení průměru sítě
- nižší režie
 - méně přenášených dat
 - větší interval mezi výměnou dat (při změně, jinak každých 30 minut)

- Border Gateway Protocol
 - v. 4 (RFC1771)
 - aktuálně RFC4271
- převládající externí routovací protokol (Exterior Gateway Protocol – EGP)
- zajišťuje směrování mezi AS
 - respektuje politiku – pravidla a dohody jednotlivých AS, resp. providerů (ISP), např.
 - např. oddělení akademického a komerčního provozu
 - cena za přenos dat
- exchange points (např. NIX)



Děkuji za pozornost