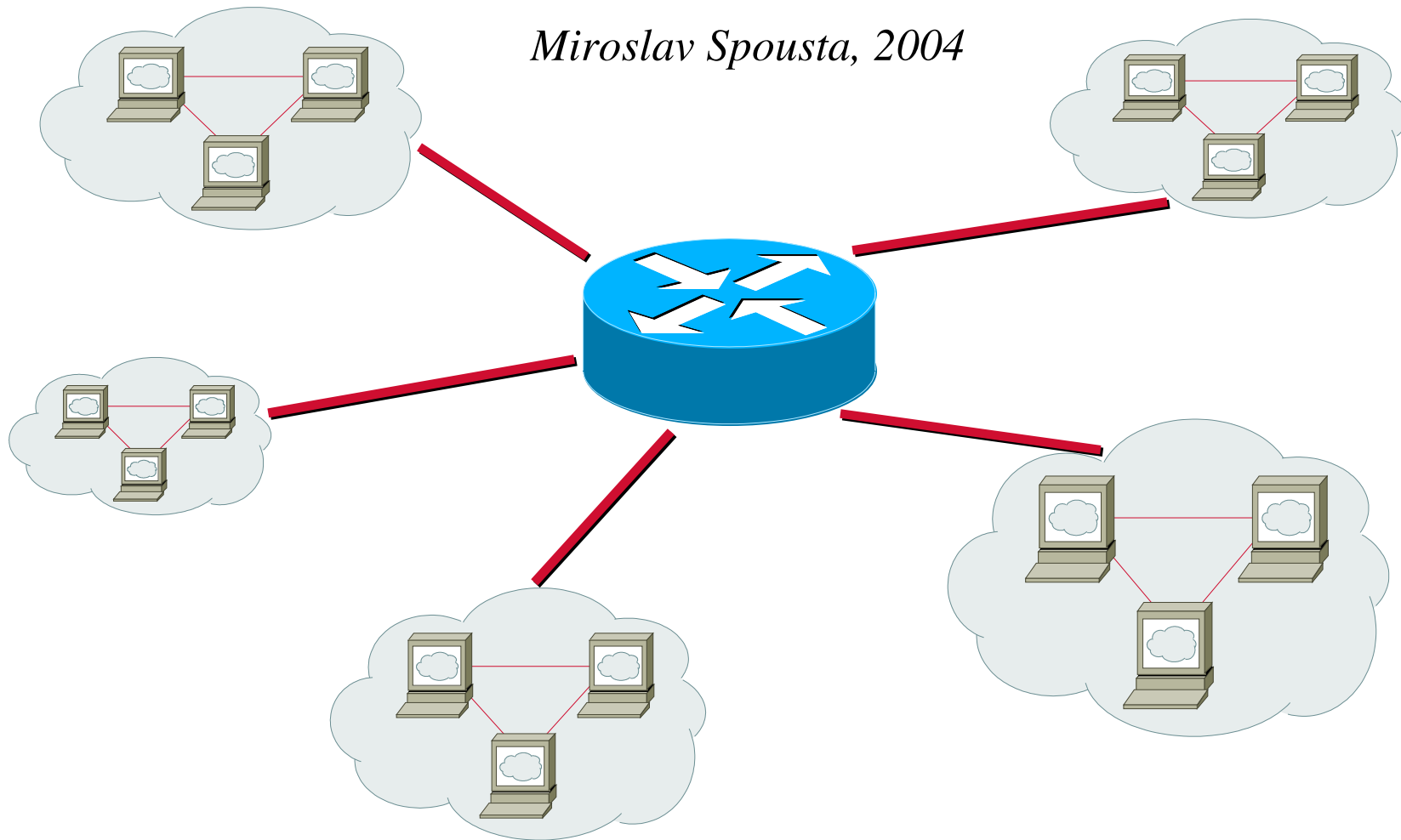


Počítačové sítě II

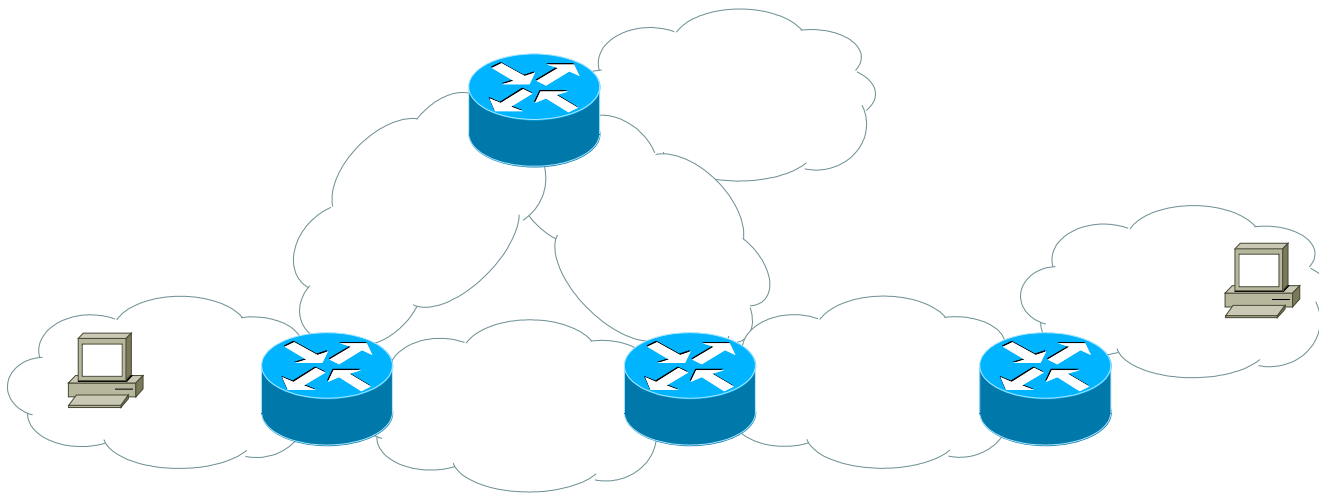
13. Směrování

Miroslav Spousta, 2004



Představa propojení sítí

- sítě jsou propojeny pomocí směrovačů
- mezi každými dvěma uzly existuje cesta přes mezilehlé sítě a směrovače
 - většinou více různých cest
- je potřeba vybrat, kudy bude směřován provoz po síti
 - pro každé dva uzly v Internetu to může být jinak

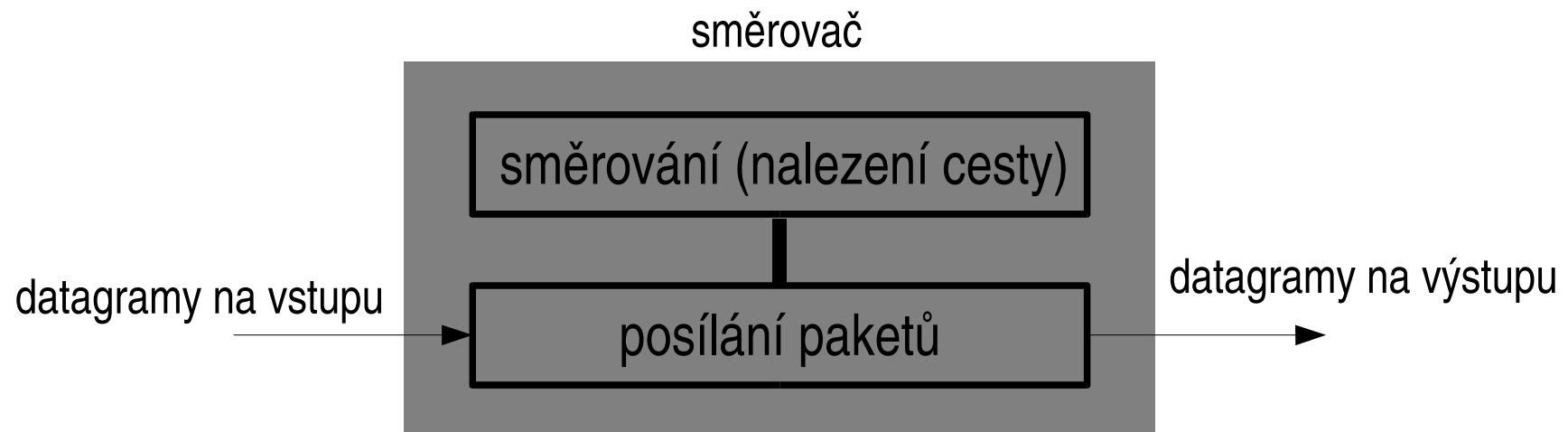


Směrování

- neboli volba směru pro další předání datagramu
- obsahuje několik podúkolů:
- výpočet optimální cesty
 - algoritmus pro výpočet nejkratší cesty v grafu
- uchování informací o cestách v grafu
 - směrovací tabulka
- předávání paketů
- udržování směrovacích informací
 - distribuce a aktualizace směrovacích tabulek
- existují různé přístupy k řešení jednotlivých částí

Směrovač

- zařízení, které pracuje na síťové vrstvě
- spojuje sítě, je připojený do několika sítí
- přijímá datagramy, rozhoduje o dalším směru datagramu sítí
- předává datagramy dalším směrovačům nebo posílá data koncové stanici v přímo připojené lokální síti



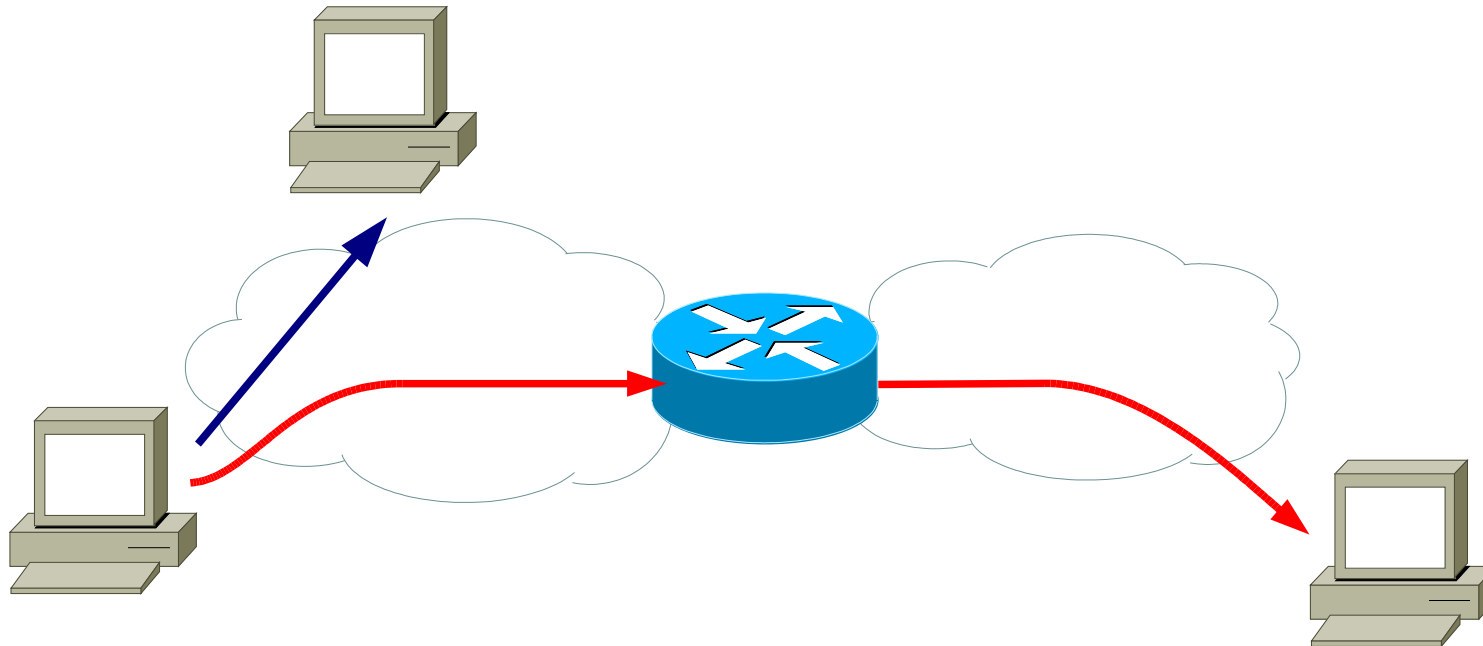
Přímé a nepřímé směrování

- Přímé směrování

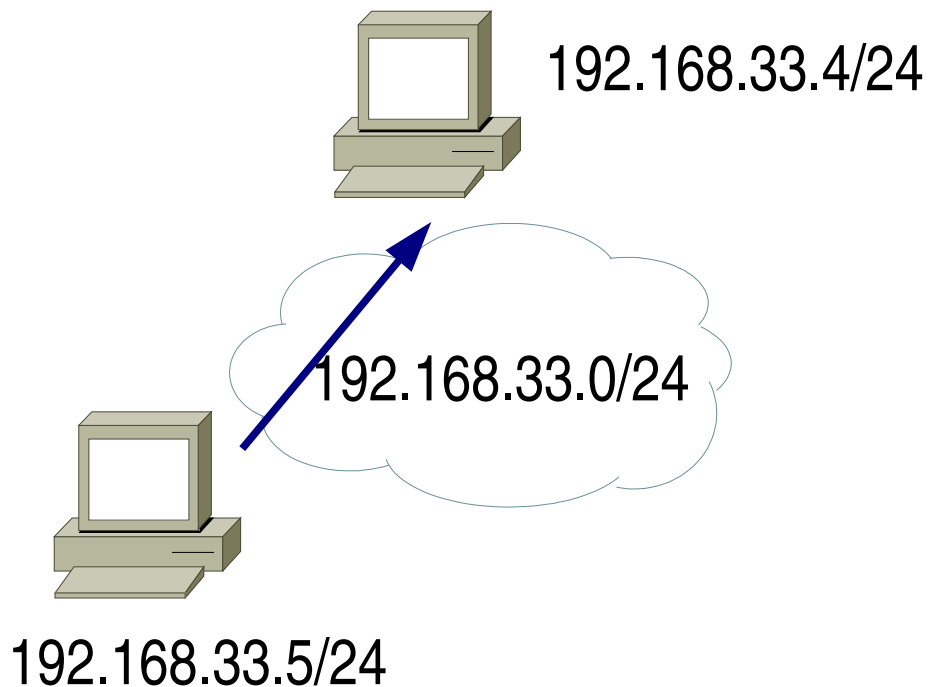
- do přímo připojené IP sítě (má stejnou síťovou adresu)
- nic se nesměruje – přímo se vyšle do lokální sítě

- Nepřímé směrování

- síťová adresa je různá od všech sítí rozhraní směrovače
- je potřeba najít v síti další směrovač, který je blíž cílové síti



Přímé směrování

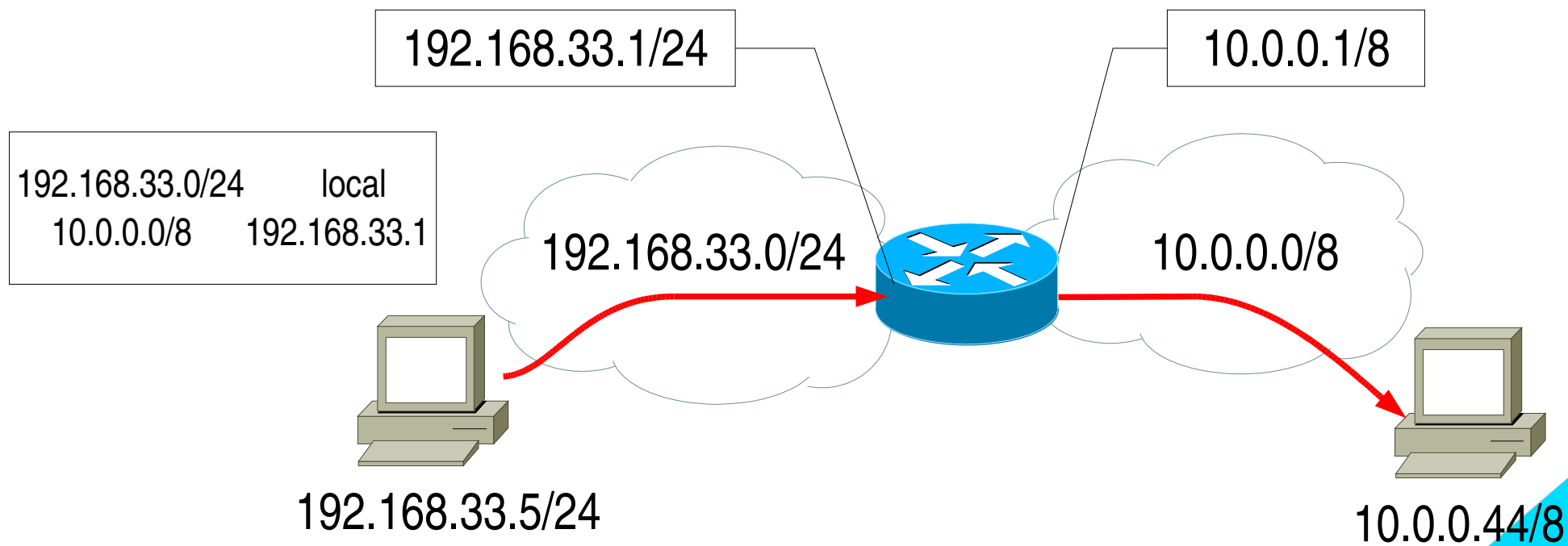


- vysílající uzel rozdělí cílovou adresu
 - podle masky vlastní sítě
- zjistí, jestli cílový uzel je ve stejné síti (má stejnou síťovou adresu, jako vysílající uzel)
- pokud ano, zjistí pomocí ARP protokolu cílovou MAC adresu
- pošle rámeček na zjištěnou MAC adresu

$$\begin{array}{|c|} \hline 192.168.33.4 \\ \hline \end{array} \text{ AND } \begin{array}{|c|} \hline 255.255.255.0 \\ \hline \end{array} = \begin{array}{|c|} \hline 192.168.33 \\ \hline \end{array} + \begin{array}{|c|} \hline 4 \\ \hline \end{array}$$

Nepřímé směrování

- pokud uzel zjistí, že adresa není lokální (má jinou síťovou složku, než všechny sítě na rozhraních stanice), je nutné předat IP datagram směrovači
- kterému, to rozhodne směrovací tabulka
- pomocí přímého směrování odešle datagram směrovači



Statické a dynamické směrování

- Statické směrování

- směrovací tabulky se nemění (jsou statické)
- je nutné nastavit směrovací tabulky ručně
- imunní vůči změnám v síti
- administrativně náročné (musí se udržovat, náchylné na chyby)
- vhodné pro malé sítě se stálou topologií

- implicitní cesta

- cesty, které nejsou jinak inzerovány (skrz firewall)

- Dynamické směrování

- obsah směrovacích tabulek se mění bez zásahu správce
- většinou je základ tabulky nastaven staticky (implicitní cesta)
- cesty do ostatních sítí se aktualizují pomocí speciálních protokolů

- dva základní druhy protokolů

- vector distance
- link state

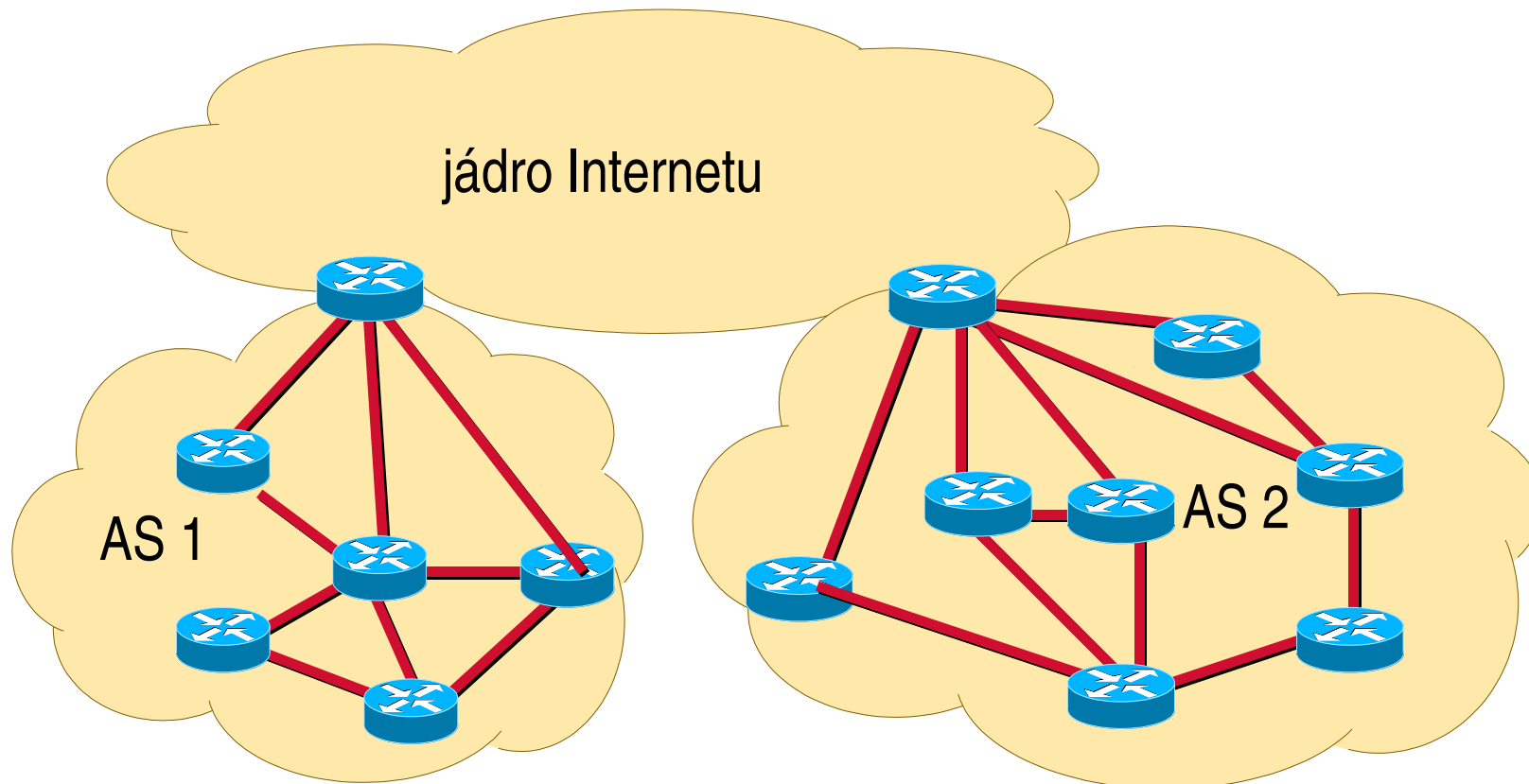
Směrovací tabulka

- základní data směrovače
- ve směrovací tabulce je:
 - cílová síťová adresa
 - metrika
 - adresa nejbližšího směrovače směrem k síti (musí být dostupný přes síť jednoho z rozhraní směrovače)
- koncový uzel má směrovací tabulku stejně jako směrovač, ale neúčastní se výměny směrovacích informací
 - je pasivní
 - většinou zná svoji síť a implicitní cestu
- jak aktualizovat směrovací informace?
 - původně v Internetu měly Core Gateways úplnou informaci o sítích
 - to bylo neúnosné => vznik autonomních systémů

Algoritmus směrování

- vezmi cílovou IP adresu
- projdi směrovací tabulku od nejspecifických k méně specifickým
- pokud záznam vyhovuje, podíváme se, zda je síť lokální (přímo připojená)
 - ano, datagram pošleme na MAC adresu příjemce
 - ne, datagram pošleme na MAC adresu sousedního směrovače
- pokud se dostaneme na konec tabulky, může tam být implicitní cesta (default route) – použijeme ji
- nenašli jsme cestu sítí => vygenerujeme ICMP zprávu Destination Unreachable

Autonomní systémy



Autonomní systém

- neinzeruje ven detailní informace
- pouze sdružené sítě
 - typicky je to informace „sítě od této IP adresy po tuto IP adresu“
- jak směruje a udržuje směrovací informace v rámci AS je na správci
 - „směrovací politika“
- původně musela být hierarchie AS přísně stromovitá
- dnes v podstatě libovolná
 - takže můžeme mít peering

EGP

- Exterior Gateway Protocols

- původně (v počátcích Internetu) se používal Exterior Gateway Protocol
- vyžadoval stromovou strukturu AS
- a jednu páteřní síť Internetu

- dnes se používá BGP (Border Gateway Protocol)

- umožňuje obecné zapojení AS (už ne pouze do stromu)
- podporuje CIDR
- několik verzí (dnes se používá verze 4)
- umožňuje nastavení priorit pro jednotlivé AS
 - např. na základě rychlosti linek, vytížení, smluv, ...

IGP

- Interior Gateway Protocols
 - protokol používaný v rámci jednoho AS
 - obecné označení pro různé protokoly směrování v rámci jednoho AS
 - politika směrování (aktualizace směrovacích tabulek) je jednotná v rámci AS
- v IP se používají hlavně dva protokoly: RIP a OSPF
- RIP (Router Information Protocol)
 - vyvinut firmou Xerox
 - používá algoritmus vector-distance (poprvé použit už v ARPANETu, 1969)
 - vhodné jen pro malé a střední sítě
- OSPF (Open Shortest Path First)
 - používá princip link state
 - používá se i ve větších sítích

IGP

- Interior Gateway Protocols
 - protokol používaný v rámci jednoho AS
 - obecné označení pro různé protokoly směrování v rámci jednoho AS
 - politika směrování (aktualizace směrovacích tabulek) je jednotná v rámci AS
- v IP se používají hlavně dva protokoly: RIP a OSPF
- RIP (Router Information Protocol)
 - vyvinut firmou Xerox
 - používá algoritmus vector-distance (poprvé použit už v ARPANETu, 1969)
 - vhodné jen pro malé a střední sítě
- OSPF (Open Shortest Path First)
 - používá princip link state
 - používá se i ve větších sítích

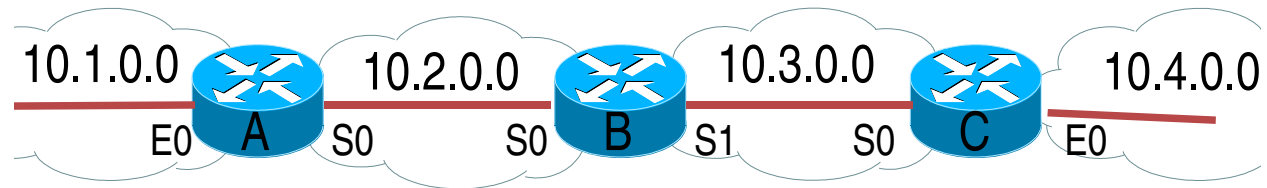
RIP

- firma Xerox v PARC, 1981
- RFC 2452 (RIPv2)
- populární protokol, implementován v UNIXu
- na konci osmdesátých let v podstatě norma pro směrování v IP sítích
- distribuovaný algoritmus – podílí se na něm všechny směrovače v síti
 - konvergence = doba, než se ustálí informace ve směrovacích tabulkách
- typu vector-distance
 - směrovače si předávají aktualizace tvořené směrovým vektorem a vzdáleností
 - vektor je vlastně adresa sítě a vzdálenost je podle metriky
- používá UDP protokol, port 520
- velmi jednoduchý na konfiguraci
- daemon – aplikační úroveň

RIP aktualizace

- ve směrovací tabulce je:
 - síťová cílová adresa
 - metrika (vzdálenost k cíli)
 - adresa směrovače (musí být dostupný přes jednu ze sítí)
 - časovač (doba od poslední aktualizace záznamu)
- RIP jako metriku používá počet směrovačů na cestě k cíli (hops)
 - maximum je 15, nekonečno je 16
 - tedy je možné mít v jednom AS maximálně 15 směrovačů
- směrovač vysílá svoji tabulku sousedům každých 30 sekund
- podle příchozích informací si upravuje svoji vlastní směrovací tabulku
 - k informacím od sousedů přičte jedničku a uloží si ji do své tabulky

RIP



10.1.0.0	E0	0
10.2.0.0	S0	0

10.2.0.0	S0	0
10.3.0.0	S1	0

10.3.0.0	S0	0
10.4.0.0	E0	0

10.1.0.0	E0	0
10.2.0.0	S0	0
10.3.0.0	S0	1

10.2.0.0	S0	0
10.3.0.0	S1	0
10.1.0.0	S0	1
10.4.0.0	S1	1

10.3.0.0	S0	0
10.4.0.0	E0	0
10.2.0.0	S0	1

10.1.0.0	E0	0
10.2.0.0	S0	0
10.3.0.0	S0	1
10.4.0.0	S0	2

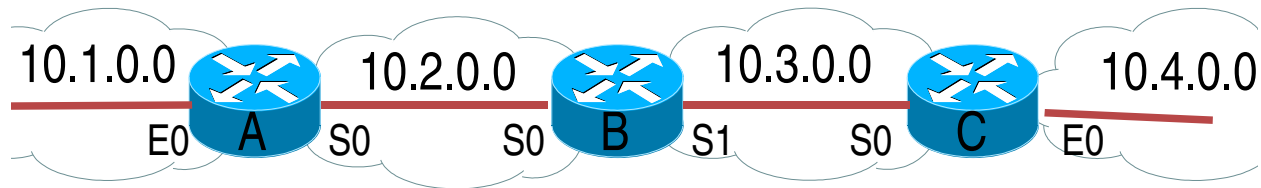
10.2.0.0	S0	0
10.3.0.0	S1	0
10.1.0.0	S0	1
10.4.0.0	S1	1

10.3.0.0	S0	0
10.4.0.0	E0	0
10.2.0.0	S0	1
10.1.0.0	S0	2

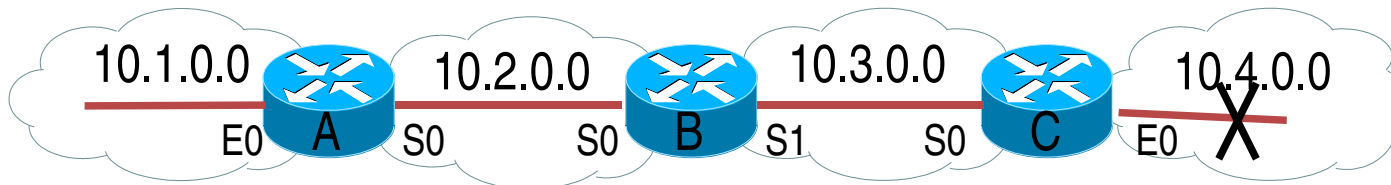
RIP konvergence

- časovač se používá pro zneplatnění neaktuálních informací
 - pokud do 180 s nepřijde od souseda informace, nastaví se v tabulce na nekonečno
 - pokud nepřijde do dalších 120 s, informace se z tabulky odstraní (garbage collection)
- RIP má problém s pomalou konvergencí
 - změna v síti se z jednoho konce sítě na druhý může propagovat dlouho
 - maximálně $15 \cdot 30s = 450s$, cca 7,5 min.
 - triggered update: při změně v topologii se propagují co nejdříve
- problém se smyčkami ve směrovacích tabulkách
 - zacyklení uživatelských paketů
 - vzniká pomalou konvergencí, způsobuje vážné problémy

RIP smyčky

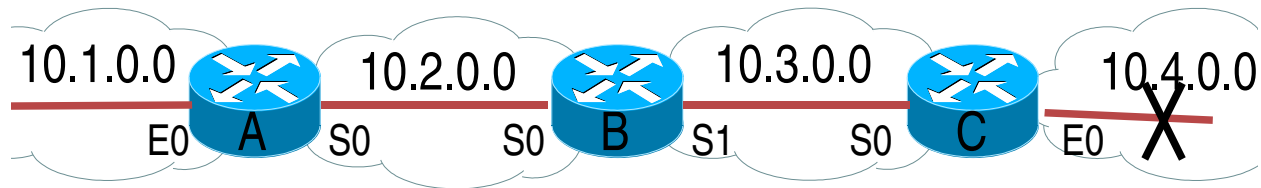


10.1.0.0	E0	0	10.2.0.0	S0	0	10.3.0.0	S0	0
10.2.0.0	S0	0	10.3.0.0	S1	0	10.4.0.0	E0	0
10.3.0.0	S0	1	10.1.0.0	S0	1	10.2.0.0	S0	1
10.4.0.0	S0	2	10.4.0.0	S1	1	10.1.0.0	S0	2

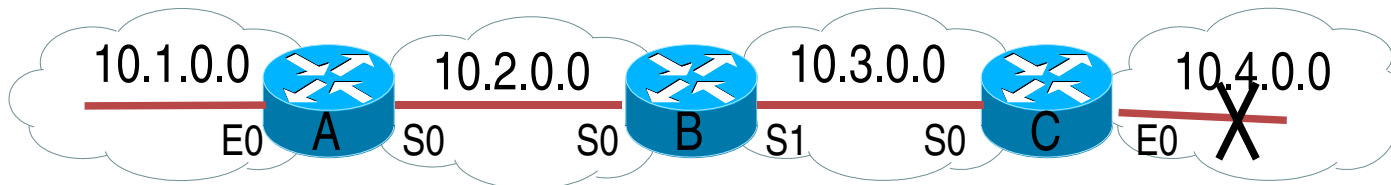


10.1.0.0	E0	0	10.2.0.0	S0	0	10.3.0.0	S0	0
10.2.0.0	S0	0	10.3.0.0	S1	0	10.4.0.0	E0	∞
10.3.0.0	S0	1	10.1.0.0	S0	1	10.2.0.0	S0	1
10.4.0.0	S0	2	10.4.0.0	S1	1	10.1.0.0	S0	2

RIP smyčky



10.1.0.0	E0	0	10.2.0.0	S0	0	10.3.0.0	S0	0
10.2.0.0	S0	0	10.3.0.0	S1	0	10.4.0.0	S0	2
10.3.0.0	S0	1	10.1.0.0	S0	1	10.2.0.0	S0	1
10.4.0.0	S0	2	10.4.0.0	S1	1	10.1.0.0	S0	2



10.1.0.0	E0	0	10.2.0.0	S0	0	10.3.0.0	S0	0
10.2.0.0	S0	0	10.3.0.0	S1	0	10.4.0.0	S0	2
10.3.0.0	S0	1	10.1.0.0	S0	1	10.2.0.0	S0	1
10.4.0.0	S0	4	10.4.0.0	S1	3	10.1.0.0	S0	2

RIP smyčka

- omezení zvyšování metriky do nekonečna
 - RIP zavedl explicitní omezení na 16
- split horizon
 - směrovač neposílá do daného síťového rozhraní informace o cestách, které mu přišly z tohoto rozhraní
- poison reverse
 - zpět po stejném rozhraní posílá metriku nekonečno (16)
- hold down
 - směrovač po obdržení informace o nedostupnosti sítě po určitou dobu nebere v potaz další informace o změně směru a metriky pro danou síť
 - minimalizace dopadu směrovacích smyček
- triggered update

RIPv1

8b	8b	16b
příkaz	verze (1)	0
typ adresy (IP = 2)		0
IP adresa		
0		
0		
metrika (vzdálenost k cíli)		

- příkaz: výzva/směrovací informace
- IP adresa: adresa sítě
- metrika: počet hopů k síti

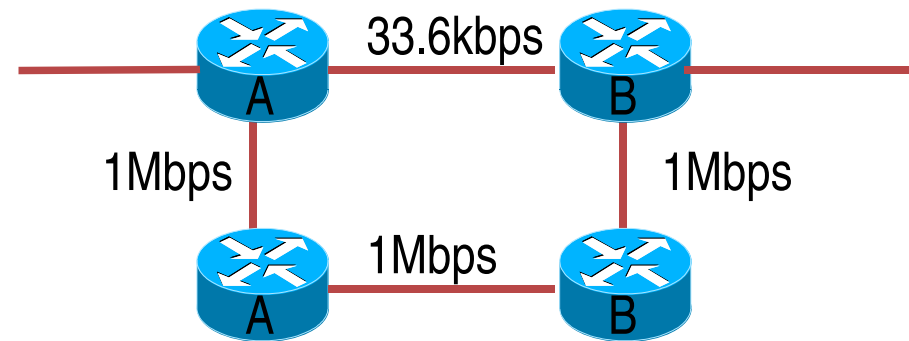
RIPv2

8b	8b	16b
příkaz	verze (2)	0
typ adresy (IP = 2)		označení cesty (tag)
IP adresa		
síťová maska		
adresa sousedního směrovače		
metrika (vzdálenost k cíli)		

- podpora CIDR
- vysílání skupinovou (multicast) adresu 224.0.0.9
- autentizace informací

Problémy RIP

- Omezená nejdelší možná cesta
 - na 15 hopů
- počítání do nekonečna (16)
- fixní metrika
 - jen hopy, žádná jiná možnost
- nelze využít paralelních cest pro rozkládání zátěže
- RIPng: pro IPv6 (RFC 2080)
 - v podstatě pouze změna IP adres
 - zůstávají zachovány ostatní vlastnosti (a problémy)

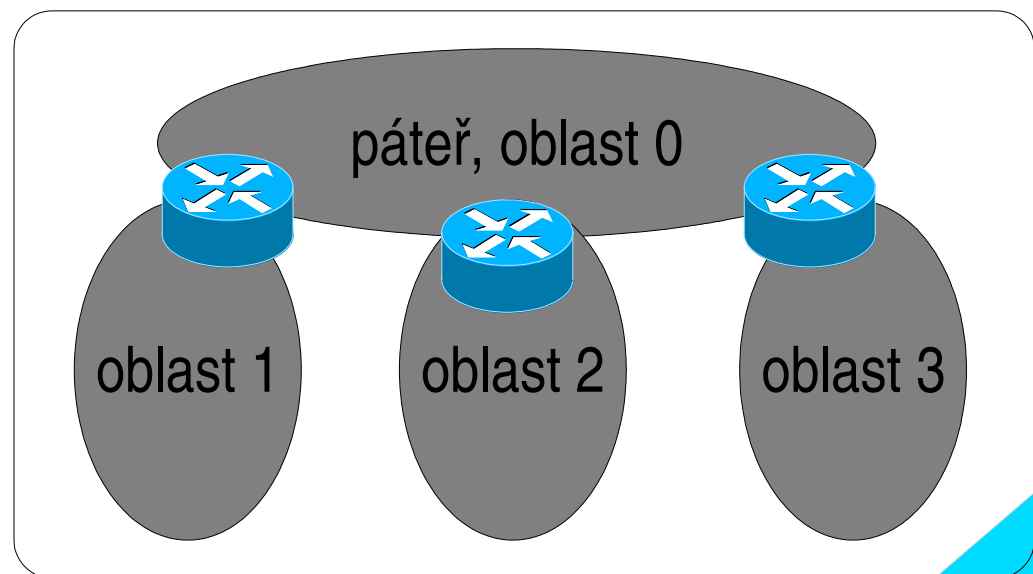


OSPF

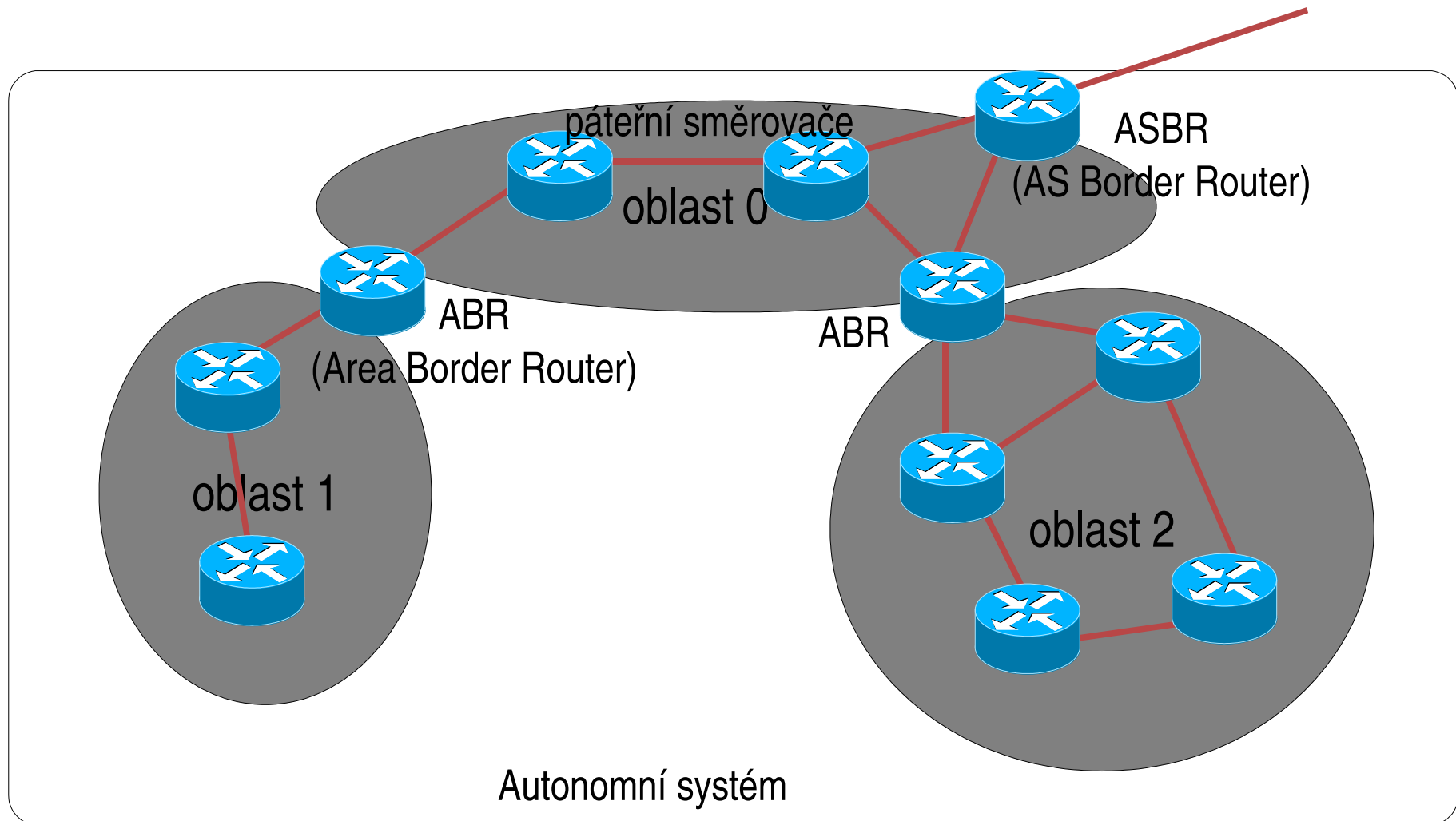
- Open Shortest Path First
 - vychází ze staršího SPF
- RFC 1245 (z roku 1991), version 2: RFC 2328 (z roku 1998)
- pracuje přímo nad IP (číslo protokolu je 89)
- algoritmus typu link-state
 - aktivně testuje stav linky k sousedním směrovačům – tuto informaci posílá do celé sítě
 - tyto informace rozesílá všem ostatním směrovačům
 - každý směrovač v síti má úplnou informaci o celé topologii sítě
 - počítá optimální cesty pomocí Dijkstrova algoritmu (sám za sebe)
- podporuje alternativní cesty, loadbalancing

OSPF metrika

- metrika: virtuální cena
 - minimalizuje se (čím menší cena, tím větší šance, že se cesta zvolí)
 - určená administrativně, měla by zahrnovat všechny potřebné vlastnosti (propustnost, vytížení, ceně)
 - vztahuje se k rozhraní směrovače (=> může být asymetrická)
- umožňuje rozdělit AS na několik oblastí (area)
 - analogie k AS, ale v jeho rámci
 - informace o vnitřní struktuře se nešíří vně oblasti
 - jedna oblast (0) je páteřní
 - bez stanic, jsou k ní připojeny všechny ostatní oblasti
 - oblast má mít max. 80 směrovačů



OSPF oblasti



OSPF směrování

- SPF algoritmus se provádí zvlášť pro každou oblast
- směrování je vlastně dvouступňové: v rámci oblasti a mezi oblastmi
- topologická databáze má tvar orientovaného grafu
- směrovač si udržuje databázi sousedů (každý jinou)
 - sousedi se objeví pomocí tzv. Hello protokolu
 - udržování sousedských vztahů
 - výměna směrovacích informací
 - vysílá se každých 10s
- dále si udržuje celkovou topologii sítě (všichni by měli mít stejnou)
- směrovací tabulky pro směrování datagramů
 - mění se podle výpočtu nejkratších cest podle aktuální topologie sítě

OSPF aktualizace

- nový směrovač
 - zjistí, kdo jsou jeho sousedi
 - s každým sousedem si synchronizuje topologickou databázi
 - poté oznámí všem ostatním směrovačům navázání sousedských vztahů
- již fungující směrovač
 - kontroluje pomocí Hello paketů, zda linky k sousedům jsou OK
 - každých 10s
 - pokud nejsou změny, jednou za 30 minut vysílá informace o sousedských vztazích
 - pokud je změna, ihned informuje

RIP vs OSPF

	RIP	OSPF
posílá se	celá směrovací tabulka	informace o přímo připojených segmentech
kdy	periodicky	ihned po změně
komu	sousedům	všem směrovačům
škálovatelnost	malá	velká
náročnost na konfiguraci	malá	velká
náročnost na síť	velká	malá
problémy	smyčky, pomalá konvergence	synchronizace