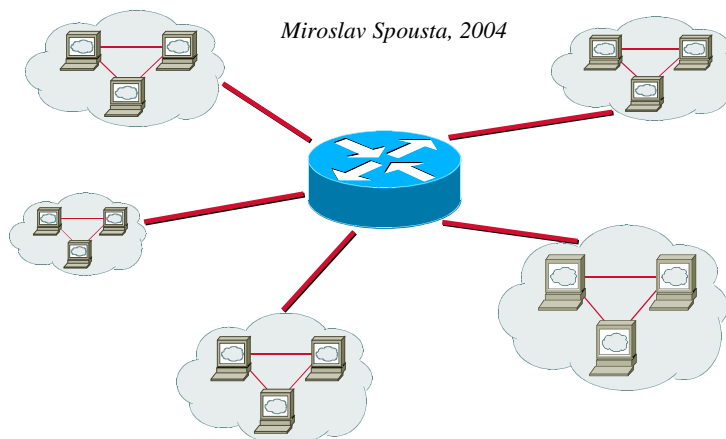


Po íta ové síť II

13. Sm rování

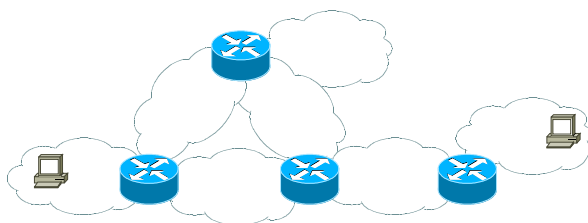
Miroslav Spousta, 2004



1

P edstava propojení síť

- síť jsou propojeny pomocí sm rova
- mezi každými dvěma uzly existuje cesta přes mezilehlé síť a sm rova e
 - v tšinou více různých cest
- je potřeba vybrat, kudy bude sm rován provoz po síti
 - pro každé dva uzly v Internetu to může být jinak



2

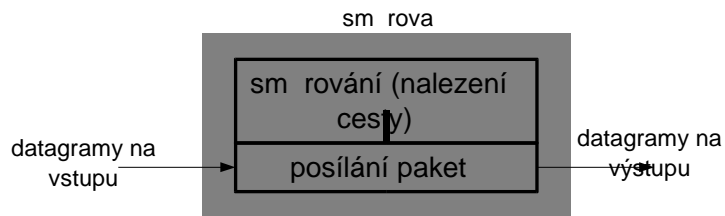
Sm rování

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

3

Sm rova

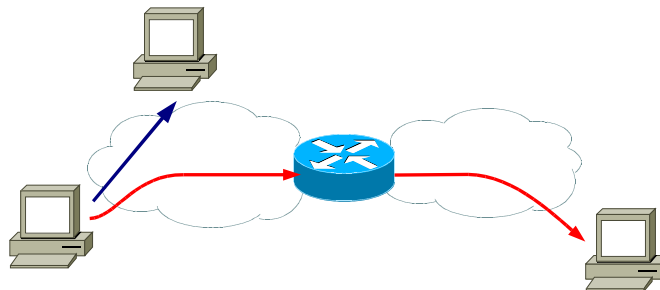
- za ízení, které pracuje na sí ové vrstv
- spojuje síť, je p ípojený do n kolika sítí
- p íjímá datagramy, rozhoduje o dalším sm ru datagramu sítí
- p edává datagramy dalším sm rova m nebo posílá data koncové stanici v p ímo p ípojené lokální síti



4

P ímé a nep ímé sm rování

- P ímé sm rování
 - do p ímo p ípojené IP síti (má stejnou sí ovou adresu)
 - nic se nesm ruje - p ímo se vyše do lokální síti
- Nep ímé sm rování
 - sí ová adresa je r zná od všech sítí rozhraní sm rova e
 - je pot eba najít v síti další sm rova , který je blíží cílové síti



5

P ímé sm rování

-
- vysílající uzel rozd ílí cílovou adresu
 - podle masky vlastní síti
 - zjistí, jestli cílový uzel je ve stejné síti (má stejnou sí ovou adresu, jako vysílající uzel)
 - pokud ano, zjistí pomocí ARP protokolu cílovou MAC adresu
 - pošle rámec na zjišt nou MAC adresu

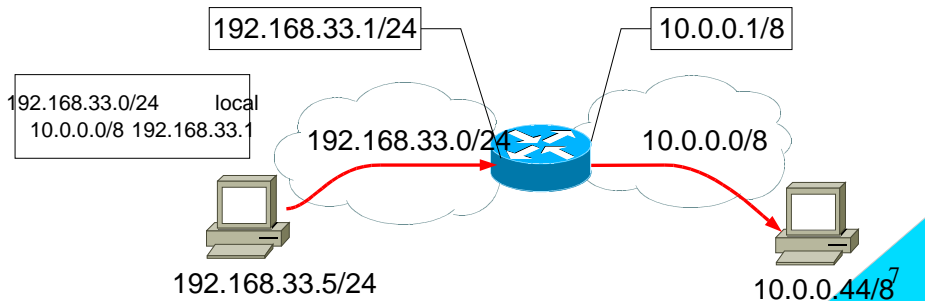
192.168.33.
4
255.255.255
.0

$$= \begin{matrix} 192.168.3 \\ 3 \end{matrix} \begin{matrix} 4 \\ + \end{matrix}$$

6

Nepřímé směrování

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



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
- **Dynamické směrování**
 - obsah směrovacích tabulek se mění bez zásahu správce
 - v tísni je základ tabulky nastaven staticky (implicitní cesta)
 - cesty do ostatních sítí se aktualizují pomocí speciálních protokolů
- **dva základní druhy protokolů**
 - vector distance
 - link state
- **implicitní cesta**
- **cesty, které nejsou jinak inzerovány (skrz firewall)**

8

Směrovací tabulka

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

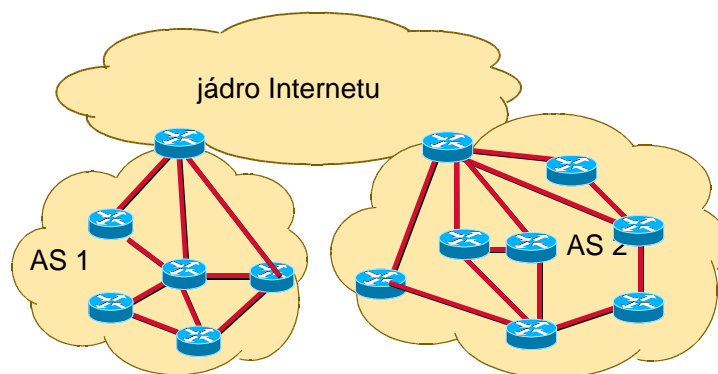
9

Algoritmus sm rování

- vezmi cílovou IP adresu
- projdi sm rovací tabulku od nejspecifických k méně specifickým
- pokud záznam vyhovuje, podíváme se, zda je síť lokální (přímě připojená)
 - ano, datagram pošleme na MAC adresu přímě
 - 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

10

Autonomní systémy



11

Autonomní systém

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

12

EGP

- Exterior Gateway Protocols
 - p vodn (v po átích Internetu) se používal Exterior Gateway Protocol
 - vyžadoval stromovou strkturu AS
 - a jednu páte ní sí Internetu
- dnes se používá BGP (Border Gateway Protocol)
 - umož ũje obecné zapojení AS (už ne pouze do stromu)
 - podporuje CIDR
 - n kolik verzí (dnes se používá verze 4)
 - umož ũje nastavení priorit pro jednotlivé AS
 - nap . na základ rychlostí linek, vytížení, smluv, ...

13

IGP

- Interior Gateway Protocols
 - protokol používaný v rámci jednoho AS
 - obecné ozna ení pro r zné protokoly sm rování v rámci jednoho AS
 - politika sm rování (aktualizace sm rovacích tabulek) je jednotná v rámci AS
- v IP se používají hlavn 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 st ední síť
- OSPF (Open Shortest Path First)
 - používá princip link state
 - používá se i ve v tších sítích

14

IGP

- Interior Gateway Protocols
 - protokol používaný v rámci jednoho AS
 - obecné ozna ení pro r zné protokoly sm rování v rámci jednoho AS
 - politika sm rování (aktualizace sm rovacích tabulek) je jednotná v rámci AS
- v IP se používají hlavn 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 st ední síť
- OSPF (Open Shortest Path First)
 - používá princip link state
 - používá se i ve v tších sítích

15

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 vlastní adresa síť a vzdálenost je podle metriky
- používá UDP protokol, port 520
- velmi jednoduchý na konfiguraci
- daemon – aplikační úroveň

16

RIP aktualizace

- ve směrovací tabulce je:
 - síťová cílová adresa
 - metrika (vzdálenost k cíli)
 - adresa směrovače (musí být dostupný přes jednu ze sítí)
 - časová (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 přichozích informací si upravuje svoji vlastní směrovací tabulku
 - k informacím od sousedů přidá jednu k sobě a uloží si ji do své tabulky

17

RIP



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

10.1.0.0 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 S1 1		

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

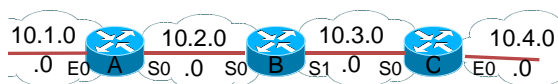
18

RIP konvergence

- asova se používá pro zneplatnění neaktuálních informací
 - pokud do 180 s nepijde od souseda informace, nastaví se v tabulce na nekonečno
 - pokud nepijde do dalších 120 s, informace se z tabulky odstraní (garbage collection)
- RIP má problém s pomalou konvergencí
 - změna v síti se z jednoho konce síť 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 v srovnávacích tabulkách
 - zacyklení uživatelských paketů
 - vzniká pomalou konvergencí, způsobuje vážné problémy

19

RIP smyčky



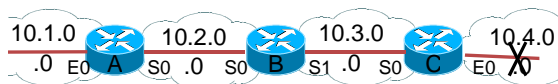
10.1.0.0 E0 0	10.2.0.0 S0 0	10.3.0.0 S0 0
10.2.0.0 S0 0	10.3.0.0 S1 0	10.4.0.0 E0 0
10.3.0.0 S0 1	10.1.0.0 S0 1	10.2.0.0 S0 1
10.4.0.0 S0 2	10.4.0.0 S1 1	10.1.0.0 S0 2



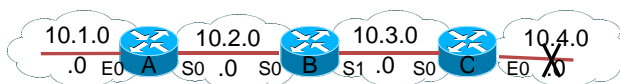
10.1.0.0 E0 0	10.2.0.0 S0 0	10.3.0.0 S0 0
10.2.0.0 S0 0	10.3.0.0 S1 0	10.4.0.0 E0 ∞
10.3.0.0 S0 1	10.1.0.0 S0 1	10.2.0.0 S0 1
10.4.0.0 S0 2	10.4.0.0 S1 1	10.1.0.0 S0 2

20

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

21

RIP slyška

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

22

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

23

RIPv2

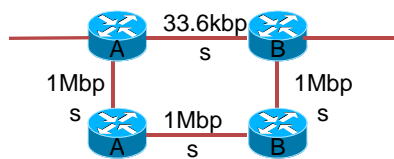
8b	8b	16b
příkaz	verze (2)	0
typ adresy (IP = 2)		označení cesty (tag)
IP adresa		
síťová maska		
adresa sousedního směrovacího rozhraní		
metrika (vzdálenost k cíli)		

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

24

Problémy RIP

- Omezená nejdelší možná cesta
 - na 15 hop
- po ítání do nekone na (16)
- fixní metrika
 - jen hopy, žádná jiná možnost
- nelze využít paralelních cest pro rozkládání zát ěže
- RIPng: pro IPv6 (RFC 2080)
 - v podstat pouze zm na IP adres
 - z stávají zachovány ostatní vlastnosti (a problémy)



25

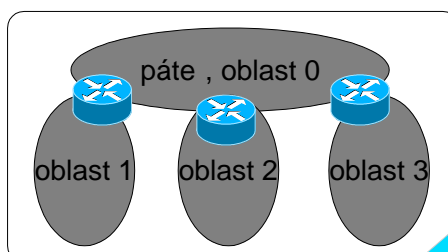
OSPF

- Open Shortest Path First
 - vychází ze staršího SPF
- RFC 1245 (z roku 1991), version 2: RFC 2328 (z roku 1998)
- pracuje p ímo nad IP (íslo protokolu je 89)
- algoritmus typu link-state
 - aktivn testuje stav linky k sousedním sm rova m – tuto informaci posílá do celé síť
 - tyto informace rozesílá všem ostatním sm rova m
 - každý sm rova v síť má úplnou informaci o celé topologii síť
 - po ítá optimální cesty pomocí Dijkstrova algoritmu (sám za sebe)
- podporuje alternativní cesty, loadbalancing

26

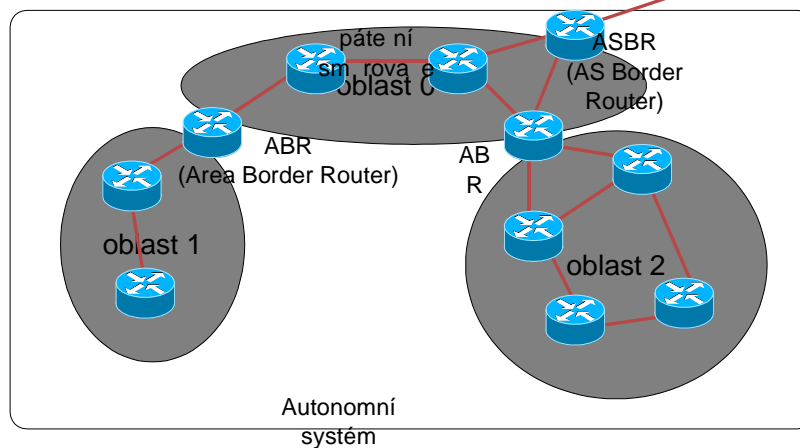
OSPF metrika

- metrika: virtuální cena
 - minimalizuje se (ím menší cena, tím v tší šance, že se cesta zvolí)
 - ur ená administrativn , m la by zahrnovat všechny pot ebné vlastnosti (propustnost, vytížení, cen)
 - vztahuje se k rozhraní sm rova e (=> m že být asymetrická)
- umož ůje rozd lit AS na n kolik oblastí (area)
 - analogie k AS, ale v jeho rámci
 - informace o vnit ní struktu e se neší í vn oblasti
 - jedna oblast (0) je páte ní
 - bez stanic, jsou k ní p ipojeny všechny ostatní oblasti
 - oblast má mít max. 80 sm rova



27

OSPF oblasti



28

OSPF sm rování

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

29

OSPF aktualizace

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

30

RIP vs OSPF

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