Směrování

-unicast RIPng a OSPF

-multicast – sms apod.

Elementární směrování

Směrovat musí umět každé zařízení podporující IPv6. Nejjednodušší přístup ke směrování se realizuje na bázi automatické konfigurace. Výchozí datovou strukturou je směrovací tabulka. Jednotlivé cíle jsou identifikovány prostřednictvím IPv6 prefixů (ukládá se jak vlastní prefix, tak jeho délka). Daný cíl může být buď přímo připojen, nebo je ve směrovací tabulce uloženo, že data směřující k cíli se mají předávat přímo připojenému směrovači na určité adrese.

***Implicitní cesta (default route)*,** slouží pro ty adresy, ke kterým v tabulce neexistuje specifický záznam. **Je dána prefixem s nulovou délkou.** Doporučuje se používat samé nuly, což je běžně zažitá konvence. Implicitní cesta je tudíž **dána prefixem ::/0**.

Tabulka může **obsahovat i individuální záznamy pro jednotlivé adresy**. V takovém případě **je prefixem celá adresa a má délku 128.** Tím se implementuje cache cílů.

Jednoduchý příklad směrovací tabulky :

	cíl	*předat na adresu*	*rozhraní*
1	::1/128	::	lo
2	fe80::215:cff:fe5d:a/128	::	lo
3	fe80::/64	::	eth0
4	2001:db8:1:3:215:cff:fe5d:a/128	::	lo
5	2001:db8:1:3::/64	::	eth0
6	ff00::/8	::	eth0
7	::/0	fe80::21b:8fff:feff:ff	eth0

Je zde jediná ethernetová karta (rozhraní *eth0*), která je zapojena do lokální sítě. Z ní vede jediná cesta ven, a sice přes směrovač s (lokální linkovou) adresou fe80::21b:8fff:feff:ff.

Záznam 1 - najdete **v každém zařízení podporujícím IPv6 – jedná se o lokální smyčku (loopback)**, tedy možnost hovořit sám se sebou.

Záznamy 2 - **adresu fe80::215:cff:fe5d:a počítač přidělil svému rozhraní *eth0*** (je odvozena z jeho ethernetové adresy 00:15:0c:5d:00:0a) a směrování pro ni řeší **předáním do lokální smyčky**.

Záznam 3 - sděluje, že **všechny ostatní lokální linkové adresy (prefix fe80::/64) se doručují přímo rozhraním *eth0*.**

Záznamy 4 a 5 - **řeší totéž pro globální individuální adresu počítače**. Ta se nachází v podsíti 2001:db8:1:3::/64 (položka 5) a její identifikátor rozhraní je samozřejmě shodný s identifikátorem v lokální linkové adrese.

Záznam 6 - **skupinově adresované datagramy mají prefix ff00::/8** a jsou **doručovány přímo do rozhraní *eth0.***

Záznam 7 - **všechny ostatní adresy mají být předávány implicitnímu směrovači** s adresou fe80::21b:8fff:feff:ff (použita jeho lokální linková adresa).

Když má daný stroj předat datagram, **vyhledá ve své směrovací tabulce všechny záznamy, jejichž cíl odpovídá cílové adrese datagramu**. **Postupuje od konkrétního k obecnému** (z kandidátů vybere ten, jehož prefix je nejdelší).

Ve srovnání s IPv4 se elementární směrování nezměnilo.

Směrovací protokoly

Ve většině případů (koncové počítače) je směrovací tabulka velmi jednoduchá. Nic složitého není potřeba. Takováto tabulka **bývá statická a může vzniknout na základě bezstavové automatické konfigurace,DHCPv6 nebo být pevně nakonfigurována**.

Směrovače bývají v komplikovanější pozici.

- **propojují větší či menší množství sítí a tomu odpovídá i rozsah** jejich směrovacích tabulek.
- **měly by reagovat na změny v topologii sítě a směrování jí dynamicky přizpůsobovat**

Používají ***směrovací protokoly***, kterými se navzájem informují o situaci a na základě těchto údajů pak upravují své tabulky. Směrovací protokol je tedy nástroj, který mimo jiné slouží k výměně informací o topologii sítě a k adaptaci směrovacích tabulek podle ní.

Pro **Internet se používá hierarchické přidělování adres a odpovídající směrování**. Síť (ať už koncová, regionální či rozlehlá síť poskytovatele připojení) dostane přidělen určitý prefix a veškeré její adresy z něj vycházejí.

Mimo danou síť pak směrovačům stačí znát dotyčný prefix a podle něj dopraví data do cílové sítě. Teprve v ní se začne posuzovat podle podrobnějších (delších) prefixů, do které konkrétní části sítě má datagram směřovat.

Internal Gateway Protocol (IGP) jsou označovány ty protokoly, které slouží ke správě směrovacích tabulek uvnitř jednoho autonomního systému. Protokoly se snaží především o to, aby rychle reagovaly na změny v síti.
V současné době existují **tři IGP protokoly použitelné pro IPv6: RIPng, IS-IS a OSPFv3**.

EGP *External Gateway Protocol (EGP)* slouží k výměně směrovacích informací mezi autonomními systémy. Tyto protokoly drží Internet pohromadě – jejich prostřednictvím se směrovače dozvídají, kudy do kterého autonomního systému a jaké prefixy jsou v něm dostupné. **Vstupní směrovač autonomního systému tudíž má ve svých tabulkách prefixy do celého Internetu a musí mít adekvátní kapacitu.**

Protokoly ze skupiny EGP usilují především o to, aby vůbec zvládly obrovské objemy informací, které musí přenášet. Proto jsou ve srovnání s IGP konzervativnější a reagují pomaleji. V současné době se **používá jediný protokol této třídy – BGP4. Pro IPv6 slouží jeho upravená verze BGP4+.**

RIPng

***Routing Information Protocol (RIP)* patří mezi internetové veterány**. Byl navržen a implementován již ve velmi raných dobách a dodnes nachází využití především v koncových sítích.

Protokol má **několik závažných omezení** (především velmi malou maximální délku cesty a pomalejší reakci na změny), je **velmi jednoduchý**. Pro malou síť, kterou chcete dynamicky směrovat, je RIP nejjednodušší cestou.

RIPng představuje reinkarnaci starého dobrého protokolu. Vychází z druhé verze RIPu, která byla definována v první polovině devadesátých let ve snaze odstranit některé nedostatky původního protokolu. Ve srovnání s **RIPv2 se RIPng liší prakticky jen v tom, že používá adresy ve formátu IPv6.** Je definován v RFC 2080: *RIPng for IPv6*.

Protokol je založený na vektoru vzdáleností (distance vector). Linky a sítě, propojující jednotlivé směrovače, mají přiřazenu určitou cenu. Má-li datagram projít jistou cestou, určí se její celková cena (vzdálenost, metrika) součtem cen linek, kterými prošel. **RIPng se snaží, aby datagramy k danému cíli vždy dorazily cestou s nejmenší celkovou cenou (metrikou).**

V půlminutových intervalech údaje ze své tabulky posílá všem sousedům. Když obdrží **tabulku od některého ze sousedů, přičte si k ní cenu linky, kterou ohlášení dorazilo, a porovná s údaji ze své tabulky.** Pokud se dozví o novém cíli, lepší cestě nebo se cesta vedoucí přes tento sousední směrovač zhoršila, upraví svou tabulku.

RIPng je určen pro menší sítě, odpovídá tomu zvolená metrika pro oceňování cest. Povolenými hodnotami jsou celá čísla v rozsahu 1–15, hodnota 16 už představuje nekonečno. Správce sítě nemá prakticky žádný prostor k tomu, aby vhodně zvolenou cenou vyjádřil průchodnost. Většinou se všechny linky oceňují hodnotou 1 a cena cesty pak vyjadřuje, kolika linkami datagram po své cestě projde (hop count). **Směrovací tabulka musí pro potřeby RIPng ke každému cíli obsahovat**:

- prefix cíle (jeho hodnotu a délku)
- metriku odpovídající celkové ceně cesty
- adresu dalšího směrovače na cestě (komu má předávat datagramy směřující k tomuto cíli)
- příznak změny (zda se v poslední době změnila)
- časovače: dobu platnosti a likvidační interval

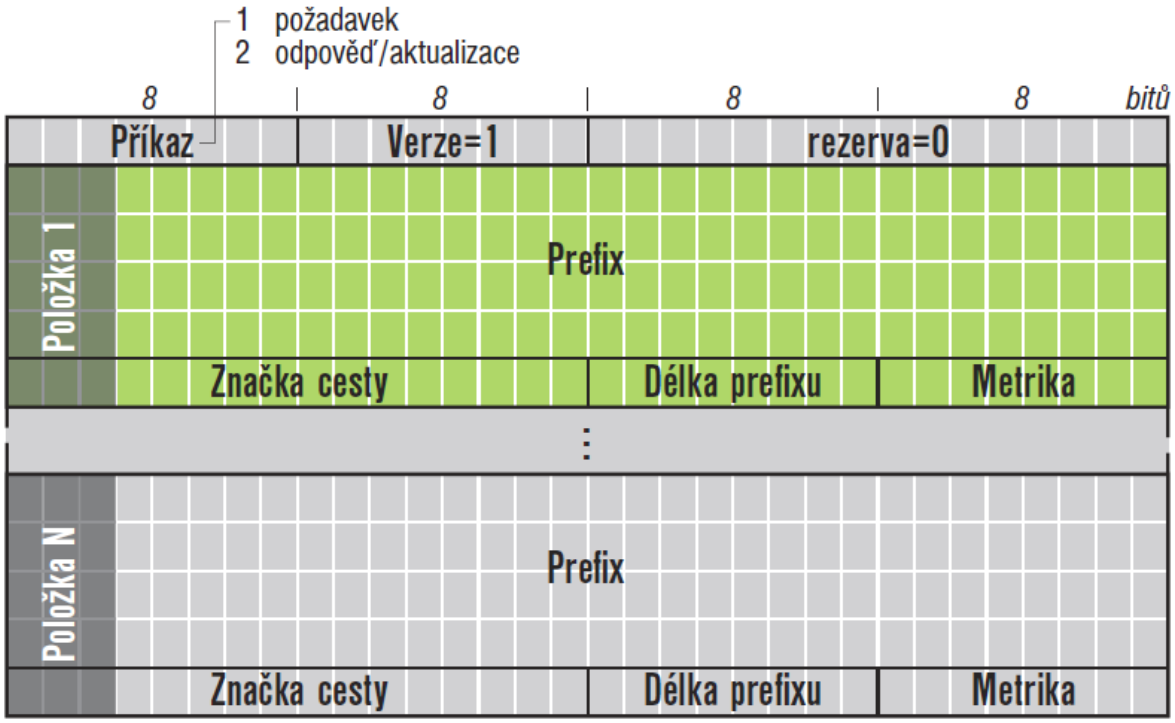
Pokud se týče sítí, ke kterým je dotyčný stroj přímo připojen, tam adresa dalšího směrovače chybí a metrika je rovna ceně této linky (zpravidla 1).

Údaje ze své tabulky posílá v následujících případech:

- *pravidelná aktualizace* zasílaná každých 30 s; aby nedocházelo k nežádoucí synchronizaci mezi zprávami jednotlivých směrovačů, je tento interval vždy posunut o náhodný čas z intervalu od –15 s do 15 s
- *aktualizace vyvolaná změnou (triggered update)*, kterou zasílá, když došlo ke změně jeho směrovacích tabulek
- *odpověď na požadavek*, když některý ze sousedů požádá o informace

V prvních dvou případech se aktualizace posílá všem sousedům. Konkrétně ji stroj zašle do všech sítí, k nimž je přímo připojen, na skupinovou adresu pro všechny RIPng směrovače ff02::9.

Formát zprávy protokolu RIPng



Příkaz (Command). Ty jsou k dispozici jen dva. Hodnota 1 signalizuje požadavek, hodnota 2 odpověď či aktualizaci.

***Verze (Version)* protokolu** - v současnosti 1

Záznamy odpovídající položkám ze směrovací tabulky.
Prefix a jeho *Délka (Prefix lenght)*, které určují cíl této položky.
Metrika (Metric), vzdálenost k cíli ze směrovače, jenž položku odeslal.
Značka cesty (Route tag) může obsahovat atribut, který je cestě přiřazen, musí s ní být uchován a dále redistribuován. Jeho prostřednictvím jsou předávány informace získané z jiných směrovacích protokolů, které v RIPng nemají žádný význam, ale při konverzi do jiného směrovacího protokolu na dalším směrovači by jej opět mohly nabýt.

Obsah zprávy

Výše uvedené tři případy rozesílání obsah zprávy tabulky se liší především tím, jaké informace bude obsahovat.

Při *pravidelné aktualizaci* se odešle kompletní směrovací tabulka.

***Aktualizace vyvolaná změnou* bude obsahovat jen ty položky, které mají nastaven příznak změny**. Následně se příznak vynuluje, protože změny byly právě ohlášeny. Cílem je, aby se sousedé o této změně dozvěděli pokud možno co nejdříve. Lze očekávat, že u nich také vyvolá změnu směrovací tabulky a následně i oni zašlou vyvolanou aktualizaci svým sousedům a tak dále. Informace se díky tomu celkem rychle rozšíří po celé síti.

Potenciální **problém spočívá v tom, aby ke změnám nedocházelo až příliš často**. Proto se vždy po odeslání zprávy s vyvolanou aktualizací nastaví čítač na náhodnou dobu z rozsahu 1–5 s. Po tuto dobu je zablokováno zasílání vyvolaných aktualizací.

***Odpověď* se váže na požadavek a obsahuje ta data, o která si vyzyvatel řekl.** Požadavek je konstruován stejně jako aktualizace, až na to, že obsahuje kód příkazu 1. Jeho položky stanoví, o které cesty má tazatel zájem. Existuje speciální případ, kdy dotazovaný má zaslat svou kompletní směrovací tabulku. Takový požadavek musí obsahovat právě jednu položku, v níž má *Prefix* i *Délka prefixu* hodnotu 0 a *Metrika* je rovna 16. Jako odpověď na tento speciální požadavek pošle dotázaný směrovač celou svou tabulku, jako při pravidelné aktualizaci.

Odpovídající projde položky v požadavku a sestaví z nich odpověď. Pokud má daný cíl ve své směrovací tabulce, vloží odpovídající informace. Jestliže cíl ve směrovací tabulce nemá, zařadí pro něj do odpovědi metriku 16. Sestavenou odpověď pak zašle žadateli.

Sekvenci požadavků a odpovědí lze využít například ke sledování stavu a činnosti směrovače. Nejčastější využití však najde při startu směrovače. Aby se co nejrychleji naučil směrovat, rozešle všem sousedům žádost o jejich kompletní tabulky a na jejich základě sestaví své.

Je-li směrovací tabulka **rozsáhlejší, může vzniknout problém s velikostí zprávy. Ta nesmí překročit MTU linky,** do které je zasílána. Pokud by měla být větší, musí se rozdělit do několika datagramů.

Jestliže má ve směrovací tabulce jako další směrovač pro daný cíl uvedeného odesilatele aktualizace, postupuje poněkud odlišně. Pokud se metrika v aktualizaci liší od směrovací tabulky, uloží si ji, nastaví příznak změny a také tentokrát požádá o odeslání vyvolané aktualizace. Je-li metrika stejná, jen si poznamená, že dotyčná cesta je stále platná.

Časovač

Každá položka má totiž přidělen **časovač, udávající její platnost**. Při jejím založení, změně či potvrzení platnosti se **nastaví na 180 s.** Dojde-li do nulové hodnoty, znamená to, že položka je neplatná a bude odstraněna.

Likvidace položky může být odstranění položky vyvolána dvěma událostmi:

- když vyprší její doba platnosti
- když směrovač, přes nějž vedla, ohlásí metriku 16.

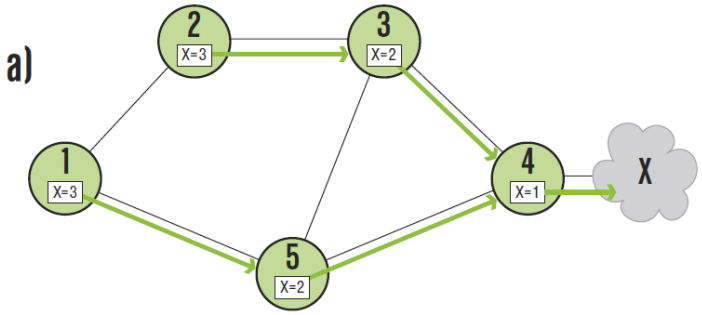
V obou případech je inicializován likvidační interval na hodnotu 120 s.

Kromě toho se položce nastaví metrika 16 a příznak změny. Výstupní části se předá požadavek na odeslání vyvolané aktualizace.

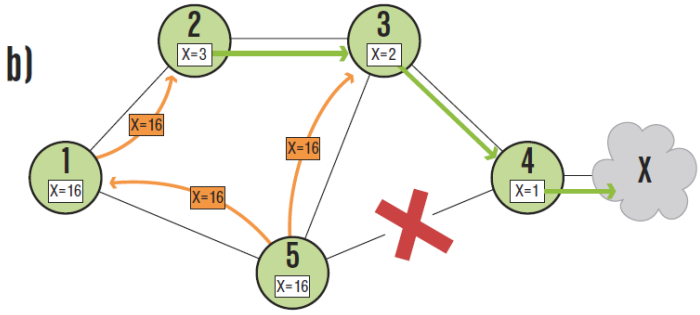
Během likvidačního intervalu je položka zařazována do všech odpovědí a aktualizací. Pokud v této době dorazí nová položka pro daný cíl, zapíší se odpovídající údaje a likvidační interval bude zrušen.

Chování RIPng

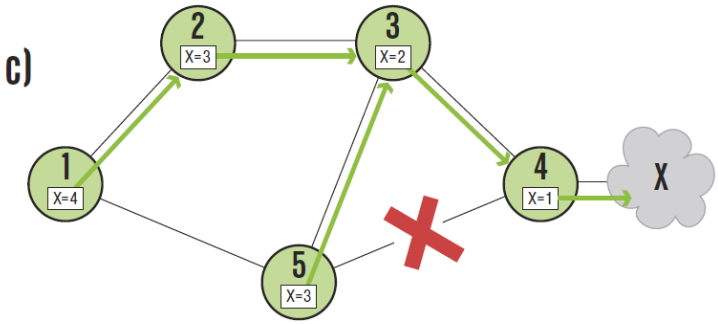
Chování RIPng v konkrétní situaci ilustruje obrázek



Výchozí situace



Přerušení spoje



Nejprve odeslal svou aktualizaci směrovač 2. Směrovač 1 se tak dozvěděl o cestě k cíli X, která vede přes 2 a má délku 4. Zanesl si ji do směrovací tabulky a ihned odeslal vyvolanou aktualizaci. Díky ní si směrovač 5 poznamenal cestu k X vedoucí přes směrovač 1 s metrikou 5 a také on odeslal vyvolanou aktualizaci, která však žádnou senzaci nezpůsobila. Když k němu za chvilku dorazila aktualizace od směrovače 3, dozvěděl se o lepší cestě s délkou 3. Vyvolanou aktualizací o tom informoval směrovač 1. Byla mu ohlášena jiná cesta k cíli X, jejíž délka se shoduje s cestou v jeho směrovacích tabulkách. RIPng v takovém případě doporučuje být konzervativní a zůstat věrný cestě, která je obsažena ve směrovací tabulce, aby se nezasílalo zbytečně mnoho vyvolaných aktualizací. Jedinou výjimkou je, pokud se dotyčné položce ve směrovací tabulce krátí životnost. Hrozí-li její brzké vypršení, je doporučeno přejít na novou stejně dlouhou cestu a předejít tak hrozící krátkodobé nedostupnosti daného cíle.

Námět k zamyšlení: Jak by se vyvíjely metriky a cesty, kdyby směrovač 3 poslal svou pravidelnou aktualizaci dříve než směrovač 2?

Jeden z problémů původního rozdělený horizont RIPu byl způsobován ohlašováním cest těm směrovačům, přes něž vedly. Například směrovače 3 a 5 z obrázku „a“ by ohlašovaly směrovači 4, že znají cestu k síti X s metrikou 2, která vede právě přes směrovač 4. Kdyby ten ztratil spojení se sítí X a vzápětí dostal toto ohlášení, zaznamenal by si do směrovací tabulky, že k cíli X zná cestu například přes směrovač 5 s metrikou 3. Tím se však směrování zacyklí – datagramy směřující do sítě X si směrovače 4 a 5 budou předávat mezi sebou, dokud nevyprší jejich životnost. S dalšími aktualizacemi se metrika pro síť X bude zvětšovat (směrovač 4 nyní ohlásímetriku 3, takže 3 a 5 si svou metriku zvětší na 4 atd.), ale potrvá určitý čas, než dospěje k hodnotě 16, která podle pravdy vyjadřuje, že síť X je nedostupná.

Vyvolané aktualizace (původní RIP neměl) podstatným způsobem zkracují dobu do vyčerpání metriky. O vyloučení vytvoření smyčky se stará mechanismus nazvaný ***rozdělený horizont (split horizon) -*** cíle se neohlašují tomu, od koho je máme.

Další variantou tohoto algoritmu je ***otrávený návrat (poisoned reverse).*** V ní se dotyčné cíle do aktualizace sice zařazují, avšak přiřadí se jim metrika 16. Je doporučeno, aby směrovač používal některou z těchto dvou variant.

OSPF

Open Shortest Path First (OSPF) je podstatně mladší, složitější a rafinovanější protokol než RIP. Je založen na stavu linek (link state). To znamená, že každý směrovač v síti používající OSPF si udržuje aktuální mapu této sítě. Obsahuje informace o tom, kdo je s kým jak propojen, jaké jsou prefixy jednotlivých podsítí, ceny linek a podobně. OSPF dbá na to, aby všechny směrovače v síti měly stejnou mapu. Kdykoli u některého z nich dojde ke změně, okamžitě o tom informuje všechny ostatní směrovače, aby si aktualizovaly svou mapu. Z ní si každý spočítá strom nejkratších vzdáleností ke všem známým cílům, jehož kořenem je on sám. Tak zjistí, kudy od něj vede nejkratší cesta ke každému cíli a zanese si ji do směrovací tabulky.

Základní výhody OSPF

- velmi rychlá reakce na změny
- schopnost zajistit směrování i v rozsáhlých sítích.

V současnosti je pro IPv4 široce používána jeho druhá verze, jejíž definici najdete v RFC 2328: *OSPF Version 2*. Úpravy OSPF pro směrování IPv6 určuje RFC 5340: *OSPF for IPv6*.

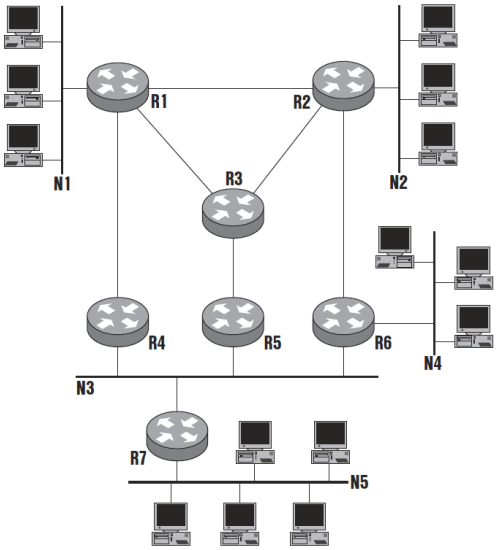
Nová verze protokolu je označována jako OSPFv3. Vedle podpory IPv6 by měla zahrnovat i různá rozšíření, navržená pro IPv4, jako je využití OSPFpro směrování skupinově adresovaných datagramů (RFC 1584) a další.

Mapa sítě (též databáze linek - LSDB) je orientovaný graf. Jeho vrcholy představují směrovače a skupinové sítě. Pro zjednodušení bereme v úvahu jen dva druhy linek:

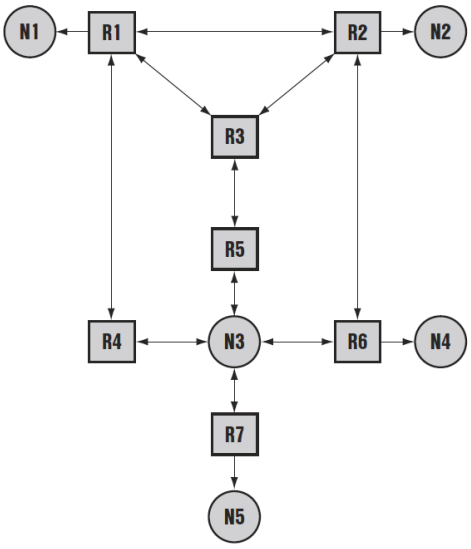
- **dvoubodové,** které navzájem spojují dvojici směrovačů
- **skupinové**, k nimž může být připojeno směrovačů několik a podporují skupinové adresování.

Typickou dvoubodovou linkou je ADSL připojení domácí sítě k Internetu.
Typickou skupinovou linkou je Ethernet.

Pokud jsou dva směrovače propojeny dvoubodovou linkou, vede mezi nimi v síťovém grafu obousměrná cesta. Jestliže cesta propojuje směrovač a skupinovou síť, dotyčný směrovač má rozhraní vedoucí do této sítě.



Příklad sítě



Reprezentace OSPF – graf – mapa sítě

Do sítí N1, N2, N4 a N5 vede pouze jednosměrná šipka. Jedná se o ***koncové sítě*** **(anglicky stub networks, čili pahýly)**. Jejich charakteristickou vlastností je, že nevedou nikam dál. Veškeré datagramy, které se v nich objeví, byly buď odeslány některým ze zdejších strojů, nebo jsou mu určeny.

N3 je příkladem ***tranzitní sítě***, kterou může procházet i provoz, který tu ani nevznikl, ani nekončí. Hrany, které ji spojují s připojenými směrovači, jsou obousměrné.

Hrany v grafu sítě jsou ohodnoceny celými čísly v rozmezí 0–65 535. Číslo představuje „délku“ příslušné cesty (oficiálně se v OSPF mluví o ceně cesty). Vzhledem ke značnému rozpětí dostupných hodnot se v ocenění hrany projevuje i rychlost linky. OSPF při hledání optimálních cest sčítá ceny linek a hledá ty, které mají nejnižší součet. Snadno tak třeba odhalí, že do daného místa bude výhodnější přepravit datagram po třech gigabitových Ethernetech než jednou ADSL linkou s rychlostí 2 Mb/s.

Sousedé a okolní

Základním prvkem OSPF je výměna informací o změnách v topologii. Vychází ze **sítě sousedů**, **kterou si protokol vytvoří**. Vznikne tak, že každý směrovač si automaticky zjišťuje, které další směrovače se nacházejí v jeho okolí. Zařadí si je do jedné ze **dvou kategorií**:

- ***Okolní směrovače (neighbors)*** jsou takové, se kterými má přímé spojení.To znamená, že je s nimi spojen dvoubodovou linkou nebo mají připojení k téže skupinové lince.

- ***Sousedé (adjacent routers)*** se vybírají z okolních směrovačů. Se sousedy si vyměňuje informace o mapě sítě (každý okolní není soused).

Jak si směrovač vybírá sousedy?

Pro dvoubodové spoje - se **navzájem propojená dvojice směrovačů vždy stane sousedy**.

Složitější situace vzniká u skupinových linek. Směrovače připojené k téže skupinové lince si ze svého středu zvolí ***pověřený směrovač (designated router)*****.** Ten se stane sousedem pro všechny zdejší stroje. Ostatní mezi sebou sousedské vztahy nenavazují.

Informace o svém okolí směrovač získá tak, že opakovaně posílá na všechna svá rozhraní zprávu oznamující jeho přítomnost – tak zvaný ***Hello* paket**.

Do něj uvede **identifikaci pověřeného směrovače pro danou síť** a také **identifikátory všech směrovačů, o nichž ví** (nedávno od nich obdržel Hello paket). Nováček se z těchto paketů dozví, kdo všechno je na dané lince přítomen a kdo je pověřeným směrovačem (pokud je).

Mapa sítě, alias databáze linek, je synchronizace map - vlastně kolekce jednotlivých oznámení o stavu v určitém místě. Oznámení se v OSPF jmenuje ***Link State Advertisement (LSA)*** a posílá je vždy ten, u nějž informace vzniká.

Typy LSA zpráv:

kód	*název*	*význam*
1	směrovač	stav rozhraní daného směrovače, posílá každý směrovač
2	síť	seznam směrovačů připojených k síti, posílá pověřený směrovač sítě
3, 4	souhrn	cesta k cíli, jenž leží mimo danou oblast (ale uvnitř AS), posílají hraniční směrovače oblasti
5	externí	cesta k cíli z jiného AS, posílají hraniční směrovače AS

Synchronizování mapy sítě

Když dva směrovače nově naváží sousedský vztah, musí si nejprve synchronizovat své mapy sítě.

Dělají to tak, že pošlou svému protějšku sadu OSPF zpráv ***Popis databáze (Database description)***, ve kterých vyjmenují identifikátory a verze LSA, tvořících jejich databázi.

Protějšek si poznamená ta LSA, která dosud neznal nebo má jejich starší verzi.

Následně o tyto LSA požádá pomocí ***Žádosti o stav linky (Link state request)*** a očekává, že mu soused pošle ***Aktualizaci stavu linky (Link state update)*** pro všechny požadované LSA - databáze jsou synchronní a oba sousedi vědí, že jejich pohled na síť je totožný.

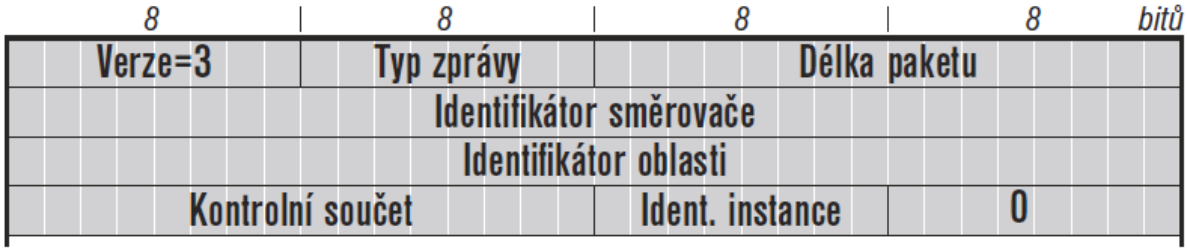
Kdykoli později dojde ke změně (například se k některému směrovači připojí nová síť), odpovídající směrovač o tom informuje všechny své sousedy prostřednictvím *Aktualizace stavu linky*. Ta obsahuje příslušné LSA, které si sousedé poznamenají do mapy a okamžitě předají dále všem svým sousedům a ti zase svým. . .

Informace se tak velmi rychle dostane do všech směrovačů v síti a mapy se opět stanou ynchronními.

Tento postup, kdy se doraziví novinka okamžitě předá všem ostatním sousedům, se nazývá **záplavový algoritmus (flooding).** Má tu příjemnou vlastnost, že nevyžaduje žádné velké přemýšlení a přitom se dostane všude, a to nejkratší cestou (protože použije všechny). Aby nedocházelo k cyklům a opakování aktualizací, předávají se dále jen ta LSA, která dotyčný dosud neznal nebo měl jejich starší verzi. Doručení aktualizace se potvrzuje (zprávou ***Potvrzení stavu linky, Link state acknowledgment*)**, aby soused věděl, že jeho oznámení dorazilo v pořádku.

Synchronizace map je velmi rychlá, je garantována a navíc se posílají jen novinky, takže OSPF má jen velmi malou režii. Kromě toho se pro jistotu stavová informace pošle ještě čas od času, i když se nezměnila. Jistota je jistota.

Hlavička OSPF zprávy



Typy OSPF zpráv

typ	*název*	*význam*
1	Hello	zjištění okolních směrovačů
2	Popis databáze	shrnuje obsah databáze
3	Žádost o stav linky	požaduje LSA
4	Aktualizace stavu linky	aktualizace databáze (posílá LSA)
5	Potvrzení stavu linky	potvrzuje aktualizaci

Hierarchie a OSPF

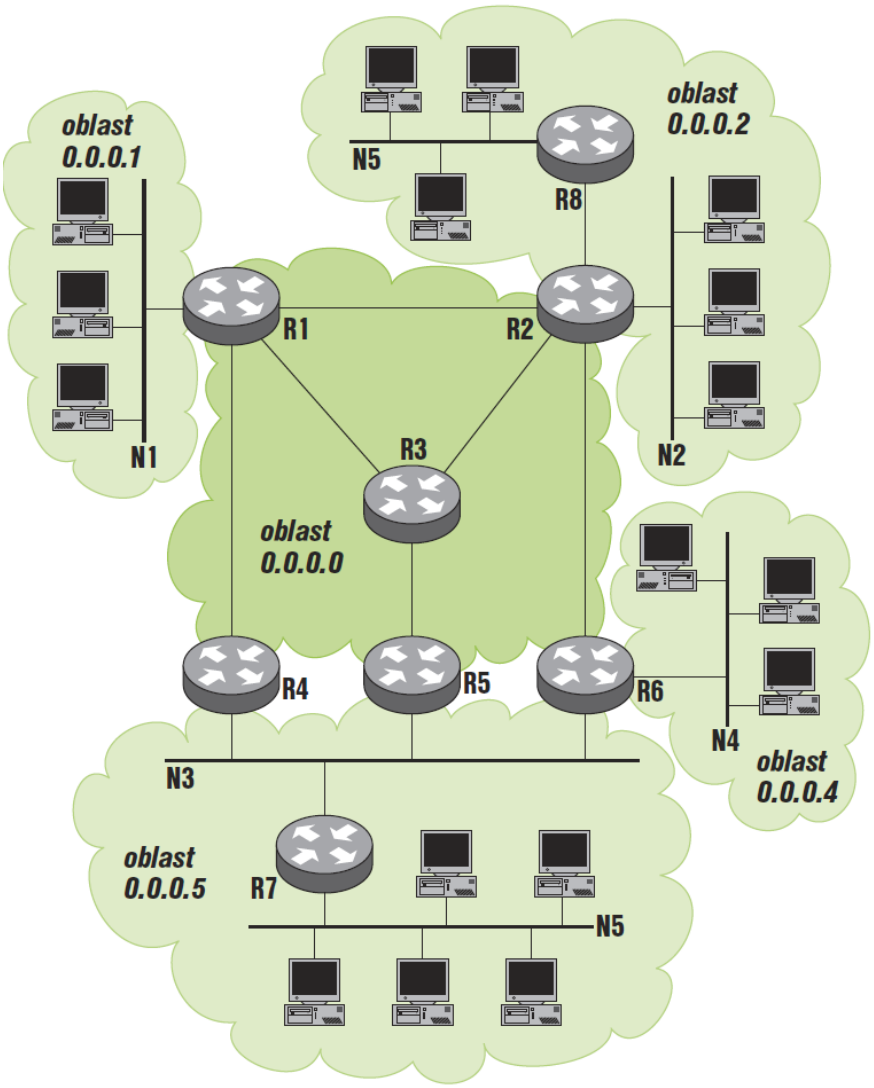
Protokol byl navržen s cílem zvládnout velké sítě, v nichž by se synchronizace map mohla stát velmi náročnou. Proto byl do OSPF zařazen koncept oblastí, které rozdělují autonomní systém na části a omezují objem přenášených směrovacích informací.

Za ***oblast (area)*** je v OSPF označována skupina souvislých sítí a strojů v nich a také všechny směrovače, které mají rozhraní do některé z těchto sítí. Každá oblast provozuje svůj nezávislý exemplář směrovacího algoritmu a udržuje si mapy, které pokrývají pouze její vlastní síť (má svojí tzv. politiku). Všechny operace, popsané výše, probíhají jen v rámci jedné oblasti. Díky tomu významně klesá režie spojená s aktualizací map.

Směrovač, který má rozhraní do více než jedné oblasti, se nazývá ***hraniční směrovač (border router).*** Musí provozovat nezávislou kopii směrovacího algoritmu a samostatnou mapu sítě pro každou z oblastí, do nichž je zapojen. V podstatě se tváří jako několik různých směrovačů – pro každou oblast jeden.

Celý autonomní systém drží pohromadě ***páteřní oblast*** **s identifikačním číslem 0.0.0.0** (identifikátory oblastí jsou 32bitové a pro jejich zápis se vžila stejná konvence jako pro IPv4 adresy). Aby bylo směrování jednoduché, **požaduje OSPF, aby všechny hraniční směrovače patřily do páteřní oblasti**. To znamená, že cestu mezi libovolnou dvojicí počítačů v různých oblastech lze rozdělit na tři části:

- První prochází oblastí s odesilatelem a končí na některém jejím hraničním směrovači.
- Druhá vede páteřní oblastí k hraničnímu směrovači, který spojuje páteřní a cílovou oblast.
- Třetí část cesty, vede cílovou oblastí od jejího hraničního směrovače k cíli.



Příklad rozdělení sítě na oblasti

Směrovač R1 má jedno rozhraní v oblasti 0.0.0.1 a tři rozhraní v páteřní oblasti 0.0.0.0. Většina směrovačů na obrázku je hraničních. Vnitřní jsou jen R3, R7 a R8, jejichž všechna rozhraní leží vždy v jediné oblasti.

Směrovače v oblasti nemusí znát detailně směrování mezi situaci za jejími hranicemi, ale oblastmi musí mít alespoň rámcový přehled o tom, jaké sítě se tam nacházejí. Vlastně jim stačí vědět, že k cíli *X* vede nejvýhodnější cesta přes zdejší hraniční směrovač *Y*.

Proto hraniční směrovače předávají do okolí informace o situaci v oblasti. Neposílají její kompletní mapu, ale pouze jakýsi souhrn. V ideálním případě shrnou celou připojenou oblast do jediného LSA záznamu, který říká třeba „za mnou leží síť s prefixem 2001:db8:abcd::/48“. Míra agregace je nastavitelná, takže si správce sítě může řídit, jak detailní informace se budou přenášet.

Souhrnné LSA záznamy jsou v ostatních oblastech šířeny stejně jako všechny ostatní. Díky nim se zdejší stroje dozvědí, jak mají směrovat datagramy určené počítačům z jiných oblastí. Například v oblasti 0.0.0.5 se díky nim všichni dozvědí, že nejlepší cesta do sítě N1 vede přes hraniční směrovač R4, zatímco do N2 a N4 to bude nejkratší přes R6. Podobně se řeší i distribuce cest k cílům z jiných autonomních systémů.

Hraniční směrovač AS, který je spojen s okolním světem, se je dozvídá prostřednictvím externího směrovacího protokolu (typicky BGP) a předává je v podobě externích LSA do páteřní oblasti. Hraniční směrovače je pak předávají dále do ostatních oblastí.

Oblast totiž lze definovat jako *koncovou (stub area)*[6] a v takovém případě se do ní externí LSA nepředávají. Směrování za hranice AS je zde prováděno pomocí **implicitní cesty**. Hraniční směrovač (nebo směrovače) propaguje do koncové oblasti implicitní cestu a říká „všechno ostatní posílejte přese mne“. Žhavými kandidáty na koncové oblasti v ukázkové síti budou 0.0.0.1, 0.0.0.2 a 0.0.0.4. Všechny mají jen jediný hraniční směrovač, takže nemají o čem přemýšlet.

Všechny výše popsané mechanismy IPv6 jsou společné pro obě verze OSPF. **Podpora IPv6 znamenala jen malé úpravy protokolu a paradoxně přinesla jeho jisté zjednodušení.**

OSPFv2 definuje bezpečnostní prvky, kterými se chrání před záškodnickými směrovači. V OSPFv3 tento prvek mizí a je nahrazen standardním IPsec přímo na úrovni IP. Došlo samozřejmě k úpravě LSA pro dlouhé adresy. Ze všech ostatních míst (jako např. identifikace směrovače či sítě) byly odstraněny IP adresy a nahrazeny 32bitovými identifikátory. Čili je tam totéž, ale jmenuje se to jinak (IPv4 add). Terminologie se změnila také u podsítí, které byly nahrazeny linkami. Sečteno a podtrženo: došlo k mírnému pokroku v mezích zákona a OSPFv3 se hodně podobá svému předchůdci.

Další část je věnovaná protokolům IS-IS a BGP4+ je nad rámec učiva. Možno pokračovat samostudiem v knize p.Satrapy.

Takže jen poznámka:
Protokol nazvaný *Intermediate system to intermediate system (IS-IS)* má v internetovém světě zcela zvláštní pozici, jedná se totiž o vetřelce zvenčí.
Původně byl vyvinut firmou Digital Equipment Corporation pro její síťovou architekturu DECnet Phase V, budiž jim země lehká. Následně jej převzala mezinárodní standardizační organizace ISO jako směrovací protokol pro svůj referenční model OSI, budiž mu země lehká.

Příjemnou vlastností je i to, že v sítích podporujících oba protokoly spravuje informace o nich pod jednou střechou. Naproti v OSPF jsou IPv4 a IPv6 dva zcela oddělené světy. V posledních letech proto některé sítě přecházejí z jiných protokolů na IS-IS. Jedním z významných příkladů je evropská akademická páteř GÉANT2.

Bez velkého přehánění lze prohlásit, že *Border Gateway Protocol (BGP)* drží pohromadě celý současný Internet. Patří mezi externí směrovací protokoly, jejichž prostřednictvím se vyměňují směrovací informace mezi různými autonomními systémy (nicméně jej lze použít i uvnitř jednoho AS). Čtvrtá verze je v současné době standardem a kdokoli chce mít svůj AS a komunikovat s okolím, musí tak činit prostřednictvím BGP4.

Skupinové vysílání čili multicast

Práce s datagramy směřujícími na skupinovou adresu se značně liší od zpracování běžných individuálních paketů**. Hlavní problém představuje směrování**. V případě skupinového vysílání se **jedná o vybudování distribučního stromu, kterým budou data šířena co nejefektivněji** ke všem příjemcům.

Doprava po Ethernetu a Wi-Fi

Nejprve, jak se skupinové datagramy dopravují po linkové vrstvě. Nejzajímavější je kombinace s Ethernetem, který má své vlastní mechanismy pro skupinové vysílání a IPv6 je využívá. Stejně je na tom i Wi-Fi, které sice používá odlišný způsob fyzického přenosu dat, ale jeho struktura adres a podpora skupinového vysílání se shodují s Ethernetem.

Ethernet podporuje skupinovou komunikaci. Jeho skupinové adresy jsou charakteristické tím, že mají **v prvním bitu jedničku, zatímco u individuálních tu najdete nulu**.

Skupinové adresy z IPv6 se do Ethernetu mapují celkem přímočaře: **vezmou se poslední čtyři bajty z cílové (skupinové) IPv6 adresy a před ně se přidá předpona 3333 (hexadecimálně).**

Např.:
Ze skupinové IPv6 adresy pro vyzývaný uzel ff02::1:ff32:5ed1 tak vznikne ethernetová adresa 33:33:ff:32:5e:d1.

Může se stát, že několik skupinových IPv6 adres splyne do jedné ethernetové – adresy, které se liší jen dosahem. Teoreticky jich je velmi mnoho, v praxi bude k těmto případům docházet jen zřídka, protože jednotlivé adresy se zpravidla liší především v závěrečném identifikátoru skupiny. Tomuto splývání se nedá zabránit.

Kdykoli začne počítač přijímat některou skupinu, musí nastavit své ethernetové rozhraní tak, aby přijímalo data přicházející na odpovídající ethernetovou adresu. Došlé datagramy dostává IPv6 vrstva a ta standardním způsobem vyřadí všechny, které dorazily na nesprávnou IPv6 adresu.

Definice přenosu skupinových IPv6 datagramů po Ethernetu je součástí RFC 2464: *Transmission of IPv6 Packets over Ethernet Networks*. Pro Wi-Fi žádné RFC neexistuje, vzhledem k podobnosti logických prvků s Ethernetem však ani není potřeba.

Multicast Listener Discovery (MLD)

Základní informaci, kterou potřebuje znát každý směrovač zapojený do skupinového života, je **seznam skupin, jež má vysílat do daného rozhraní**. V podstatě se jedná o ty skupiny, které zde mají alespoň jednoho příjemce.

Ke zjišťování příjemců sloužil ve světě **IPv4 Internet Group Management Protocol (IGMP)**. **V IPv6 má název MLD**, základní principy zůstaly podobné.

MLD se momentálně vyskytuje **ve dvou verzích.**

- **MLDv1 (RFC 2710) je protipólem IGMPv2.**
- MLDv2 odpovídá IGMPv3. Jeho definici obsahuje v RFC 3810: *Multicast Listener Discovery Version 2 (MLDv2) for IPv6*.

Novější verze protokolu umožňuje filtrovat zdroje. Zatímco v MLDv1 lze pouze ohlásit zájem o příjem skupiny G, v MLDv2 může počítač požadovat data pro skupinu G vysílaná konkrétním strojem nebo naopak příjem skupiny z určitých adres odmítnout. Obě verze spolu mohou spolupracovat, ovšem formáty jejich zpráv i používané postupy se liší.

MLD a jeho zprávy jsou jen jedním typem zpráv ICMPv6. Jejich tvar vychází ze základního formátu ICMP, který konkretizuje.

- posílají se vždy z lokální linkové adresy příslušného rozhraní
- jejich maximální počet skoků je nastaven na jedničku
- musí nést rozšiřující hlavičku *Upozornění směrovače*, aby si jich všímaly i směrovače, které samy neposlouchají cílovou skupinu příslušného datagramu.

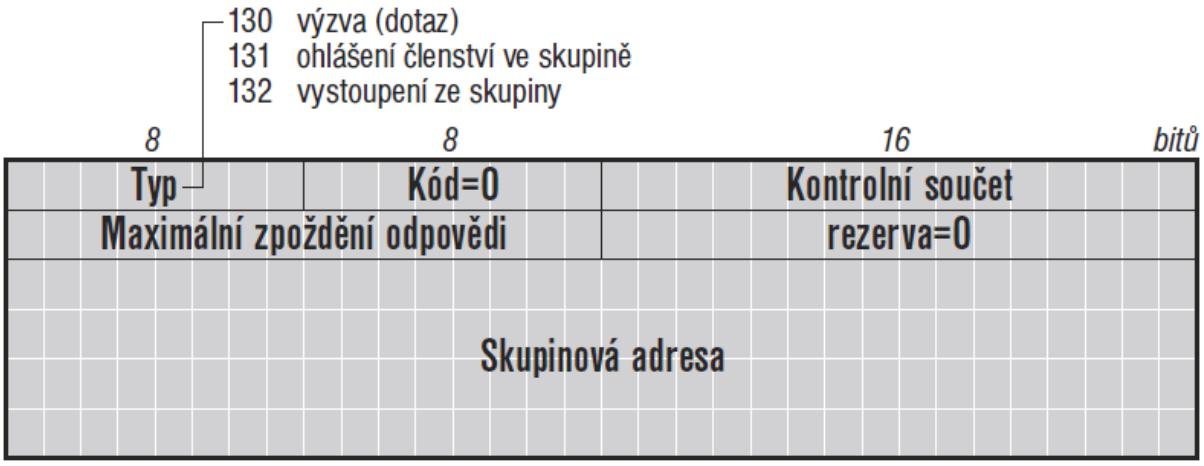
V prvním poli zprávy je v souladu s pravidly ICMPv6 vždy uveden typ. Typy i formáty zpráv se poněkud liší v závislosti na verzi protokolu.

zpráva	*MLDv1*	*MLDv2*
dotaz (query)	130	130
hlášení (report)	131	143
ukončení (done)	132	–

MLD verze 1

MLDv1 se nestará o odesilatele. Zajímá jej pouze informace, zda pro danou skupinu existuje někdo, kdo ji přijímá. Směrovač si pro každé rozhraní udržuje seznam skupinových adres, pro které se zde nachází alespoň jeden posluchač. Z těchto informací se pak vychází při stanovení distribučních stromů pro jednotlivé skupiny a směrování skupinově adresovaných datagramů. Na rozhraní, které podporuje skupinový provoz, je směrovač povinen přijímat data adresovaná kterékoli skupině – včetně těch, které sám nezná.

Absence informací o odesilatelích se odráží i v jednodušším formátu zpráv protokolu. Ten je pro všechny tři jejich typy společný



Když počítač **vstoupí do nové skupiny, pošle na její adresu MLD zprávu typu 131** ohlašující členství v této skupině. Směrovače si u příslušného rozhraní přidají skupinovou adresu do seznamu vysílaných (pokud ji tam ještě nemají). Doporučuje se poslat toto ohlášení opakovaně, aby se eliminovala případná ztráta datagramu.

Pokud počítač naopak **ukončuje svou účast ve skupině, pošle MLD zprávu typu 132,** tedy ukončení členství. Tato zpráva se neposílá na adresu skupiny, ale na ff02::2, což je skupinová adresa pro všechny směrovače na dané lince.

V případě kdy příjem skupiny naposledy ohlašoval právě ten počítač, který se odhlašuje si směrovač musí udělat jasno. **Pošle proto na adresu skupiny dotaz – MLD zprávu typu 130, v níž je jako *Skupinová adresa* (Multicast address) uvedena adresa zjišťované skupiny**.

Reakce na dotaz je definována následovně:
Každý počítač, který dostane dotaz , si nastaví časovač na náhodný interval.Jeho horní hranici udává položka *Maximální zpoždění odpovědi (Maximum response delay)* v dotazu. Po vypršení intervalu pošle na adresu skupiny ohlášení svého členství. Jakmile dostane **ohlášení od jiného člena, časovač zruší a sám své ohlášení už posílat nebude.** To znamená, že ze skupiny odpoví vždy jen jeden náhodně vybraný stroj.

Občas některý počítač (objevování skupin) přestane poslouchat skupinu, aniž by to korektně ohlásil. Proto směrovače opakovaně posílají do připojených sítí obecné dotazy. Jedná se o MLD zprávu typu 130, kde *Skupinová adresa* je nulová. **Směrovač tím říká „chtěl bych vědět, které všechny skupiny zde mají posluchače“. Dotaz zašle na adresu ff02::1 (všechny uzly na lince).**

Počítač se chová stejně, jako kdyby naráz dostal konkrétní výzvu pro všechny své skupiny. Aby se počet dotazů udržoval v rozumných mezích, posílá je vždy jen jeden ze směrovačů připojených k dané lince – ten, který má nejmenší IPadresu. Implementace jeho výběru je velmi jednoduchá. Každý směrovač poslouchá a pokud příliš dlouho nedorazí obecný dotaz od někoho s menší adresou než je jeho vlastní, pošle jej sám.

Dotazy tedy posílá jen jeden směrovač, ale odpovědi dostávají a zpracovávají všechny. Díky tomu si všechny směrovače připojené k dané lince udržují konzistentní informace o zdejších skupinách.

MLD verze 2

Druhá verze protokolu **přináší možnost omezit příjem skupinových dat v závislosti na jejich zdroji.** Dává na výběr dvě varianty, jak skupinově adresované datagramy filtrovat.

- je možné požadovat doručování skupinových dat jen od vybraných stanic. Takový režim je označován jako **INCLUDE(L), kde L představuje seznam přijímaných adres**.

- druhou možností je příjem dat od všech, kromě několika uvedených zdrojů, který se označuje **jako EXCLUDE(L)**.

Filtrování je důsledně promítnuto do celého systému.

Aplikačním rozhraní - kde si program může poručit, od koho chce či nechce dostávat skupinová data.

Síťová vrstva počítače si musí vést evidenci o požadavcích aplikací pro jednotlivé sokety a podle nich pak vytvářet stav příjmu skupinových dat na každém ze svých síťových rozhraní.

Směrovač musí kombinovat **zprávy od různých klientů a dát z nich dohromady přehled o tom, jaké skupinové datagramy předávat do sítí, k nimž je připojen**.

Jsou stanovena pravidla, jak kombinovat jednotlivé požadavky. Ty mohou obsahovat nejen různé adresy zdrojů, ale mohou používat i různé režimy filtrování.

INCLUDE(X) + INCLUDE(Y)	→	INCLUDE(X ∪ Y)
EXCLUDE(X) + INCLUDE(Y)	→	EXCLUDE(X – Y)
EXCLUDE(X) + EXCLUDE(Y)	→	EXCLUDE(X ∩ Y)
příklady	→	
INCLUDE(a, b, c) + INCLUDE(c, d, e)	→	INCLUDE(a, b, c, d, e)
EXCLUDE(a, b, c) + INCLUDE(c, d, e)	→	EXCLUDE(a, b)
EXCLUDE(a, b, c) + EXCLUDE(c, d, e)	→	EXCLUDE(c)

Jestliže se spojují **dva filtry typu INCLUDE, vznikne filtr stejného typu zahrnující sjednocení adres** původních dvou.

Při kombinaci **EXCLUDE s INCLUDE dostane přednost první z nich, protože požaduje data od všech kromě několika uvedených**. Z nich budou vyloučeny ty adresy, o které žádá filtr typu INCLUDE.

Spojením dvou filtrů typu **EXCLUDE vznikne filtr stejného typu vylučující jen ty adresy, které nechce ani jeden z původních filtrů.**

Toto spojování probíhá v několika místech. Provádí je každý příjemce, když dává dohromady požadavky jednotlivých aplikací, které pak souhrnně ohlašuje jako požadavky pro své síťové rozhraní.

Podobně pak postupuje i směrovač, který slučuje požadavky jednotlivých klientů.

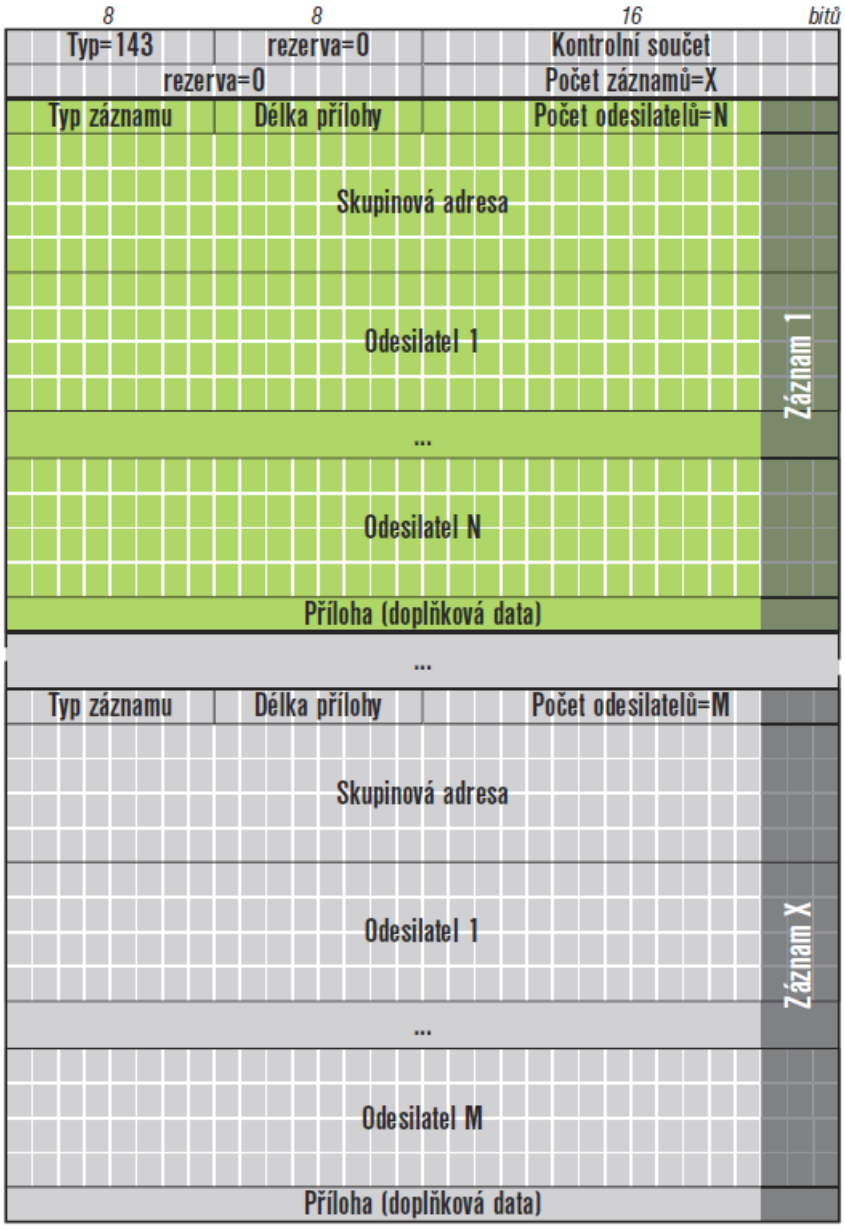
V praxi nejčastější případy

Příjem skupinové adresy ze všech zdrojů, tedy bez jakéhokoli filtrování. Tato situace je v MLDv2 vyjádřena jako **EXCLUDE( ), tedy EXCLUDE s prázdným seznamem odmítaných adres.** Jakmile se objeví při kombinování požadavků, je jasné, že i výsledkem

Druhým případem je vystoupení ze skupiny, tedy ukončení jejího příjmu. Z pohledu MLDv2 příslušný filtr přejde do režimu INCLUDE( ).

MLDv1 rozlišuje na straně příjemce skupinových změna příjmu dat dvě základní situace: vstup do skupiny a její opuštění. Naproti tomu MLDv2 má jen jednu událost tohoto typu – změnu v příjmu skupin. Zahájení či ukončení příjmu skupiny představují její speciální případy, kromě nich může změna zahrnovat i rozšíření nebo zúžení počtu přijímaných zdrojů nebo změnu režimu filtrování příslušné skupiny. Pokud dojde k jakékoli z těchto událostí, příjemce pošle MLD zprávu typu *Hlášení* na adresu ff02::16 pro všechnyMDLv2 směrovače na lince.

Formát hlášení MLDv2



Do jednoho hlášení lze v MLDv2 vložit celou řadu informací. Ty jsou obsaženy v tak zvaných záznamech a úvodní hlavička hlášení obsahuje především údaj o tom, kolik záznamů se nachází uvnitř. **Záznamy mají různý význam, který je určen položkou *Typ záznamu (Record type)***.

typ	*význam*
1	MODE_IS_INCLUDE
2	MODE_IS_EXCLUDE
3	CHANGE_TO_INCLUDE
4	CHANGE_TO_EXCLUDE
5	ALLOW_NEW_SOURCES
6	BLOCK_OLD_SOURCES

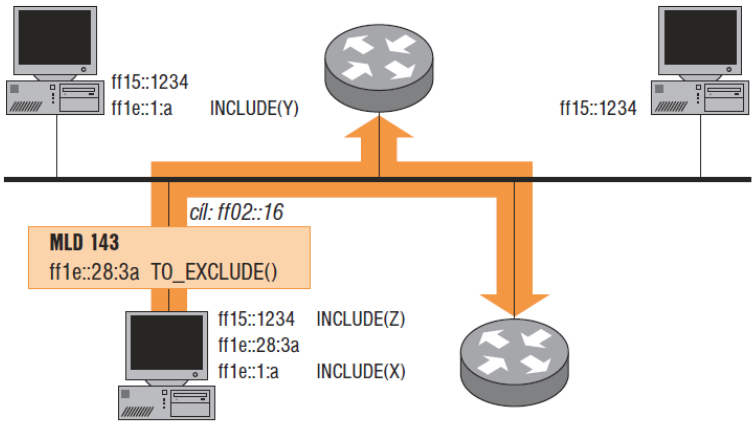
Pokud se u stanice něco změní na příjmu skupiny (či několika skupin), okamžitě odešle MLD hlášení s popisem změn.

Došlo-li ke změně režimu filtrování skupiny, pošle pro ni záznam typu 3 (při změně z EXCLUDE na INCLUDE) nebo 4 (naopak). Vloží do něj i adresy zdrojů, které nový filtr obsahuje.

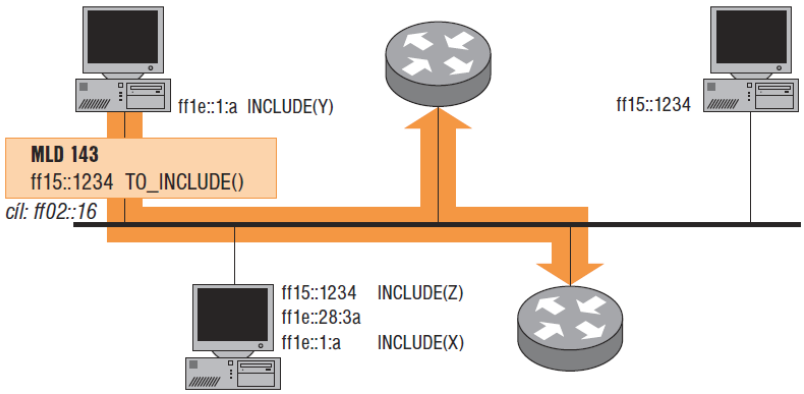
Pokud režim zůstal zachován, ale došlo ke změně filtrovaných adres, posílá stanice záznamy typu 5 (chce přijímat nové adresy) a 6 (končí příjem dříve přijímaných).

Níže uvedená tabulka shrnuje, jaké záznamy se posílají pro ohlášení změny filtru ze stavu „před“ do stavu „po“. A a B v ní reprezentují seznamy adres.

před	*po*	*odeslané záznamy*
INCLUDE(A)	INCLUDE(B)	ALLOW(B–A), BLOCK(A–B)
INCLUDE(A)	EXCLUDE(B)	TO_EXCLUDE(B)
EXCLUDE(A)	EXCLUDE(B)	ALLOW(A–B), BLOCK(B–A)
EXCLUDE(A)	INCLUDE(B)	TO_INCLUDE(B)



Počítač vstupuje do skupiny

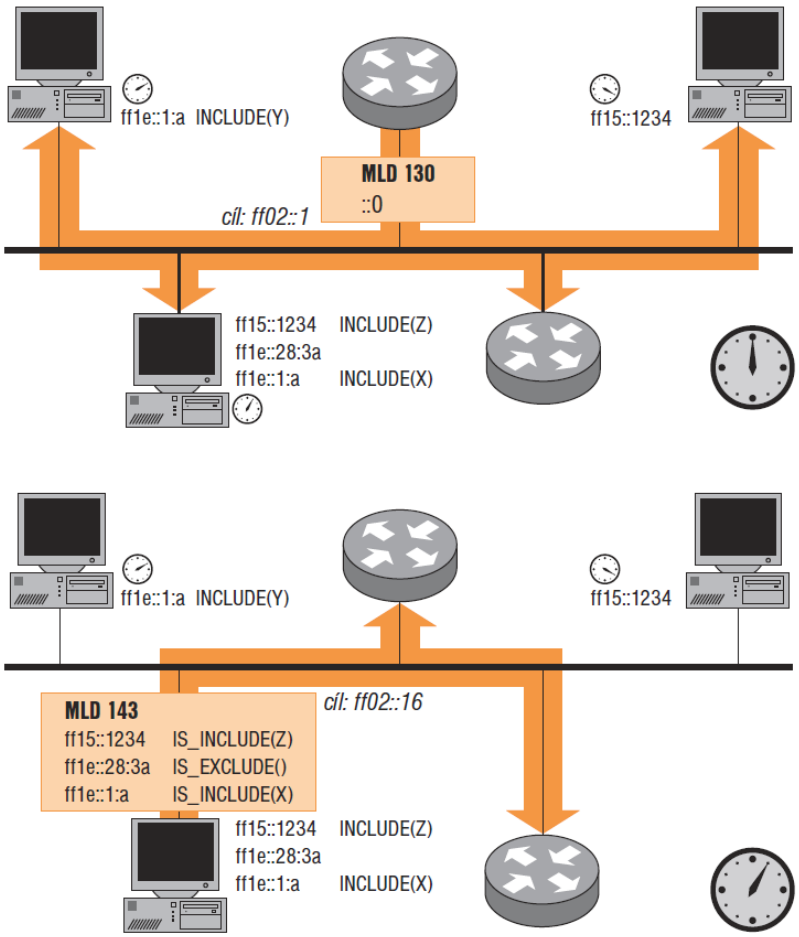


Počítač opouští skupinu

K zjištění aktuálního stavu směrovačem slouží MLD dotaz (query). V MLDv2 může směrovač podle potřeby položit jeden ze **tří typů dotazů**:

Obecný dotaz s nulovou adresou skupiny znamená „Sháním všechny posluchače skupinových dat“

Pokud uvede adresu skupiny, zajímají jej jen příjemci této skupiny. Jestliže navíc přidá i adresy zdrojů, ptá se „Poslouchá někdo tuto skupinu z uvedených zdrojů?“ Obecný dotaz se posílá všem (na adresu ff02::1), konkrétní na adresu skupiny.



Směrovač zjišťuje aktivní skupiny – následuje ještě MLD 143 dalších dvou stanic

Směrování skupinových datagramů

Protokol MLD umožňuje klientům vstoupit do skupiny a přihlásit se k odběru datagramů, směřujících na její adresu. Dále je potřeba:

- **koordinovat informace o skupinách**
- **zjistit rozložení příjemců jednotlivých skupin v síti**
- **vytvořit co nejefektivnější cesty, jak jim doručovat data.**

To zajišťují **skupinové směrovací protokoly.**

Prodělaly vývoj, v němž významnou roli hrál ***Distance Vector Multicast Routing Protocol (DVMRP)***, který ale není efektivní a nehodí se pro rutinní nasazení.

V současné době dostává přednost **skupina protokolů s názvem *Protocol Independent Multicast (PIM)*.**

Společné označení skrývá čtyři protokoly:

PIM – Dense Mode (PIM-DM) je vhodný pro situace s vysokou hustotou příjemců, kdy je třeba datagramy distribuovat skoro do všech částí sítě. Jeho použitelnost je prakticky omezena na lokální sítě a v současné době je považován za překonaný.

PIM – Sparse Mode (PIM-SM) naopak předpokládá, že příjemci jsou v síti roztroušeni jen zřídka. Vytváří pro ně distribuční stromy na základě žádostí o příjem skupinových dat. To jej činí **vhodným pro rozlehlé sítě a široko daleko nejpoužívanějším protokolem současnosti.**

Bidirectional PIM (BIDIR-PIM) představuje variantu PIM-SM. Jeho distribuční stromy jsou obousměrné, zatímco PIM-SM používá skupinu jednosměrných stromů.

PIM Source Specific Multicast (PIM-SSM) rozlišuje při doručování nejen skupinovou adresu, ale i adresu zdroje. Je určen především pro komunikaci, kdy vysílá jediný zdroj, jehož data přijímá řada klientů (něco jako internetová televize či rádio).

K některým svým činnostem skupinové směrování využívá i směrovací informace pro individuální (unicastové) pakety. Nestará se však o to, jak individuální směrovací tabulka vznikla a není vázáno na konkrétní směrovací protokol. Odtud pochází ono „Protocol Independent“ v názvu PIM.

Při skupinovém směrování je opět hierarchicky uspořádaná struktura.

PIM domény - V Internetu najdeme oblasti, uvnitř nichž běží vnitřní směrovací protokol (zpravidla PIM-SM). PIM domény, mají svá shromaždiště (RP) a organizují si doručování skupinových dat po svém. **Představují určitou analogii autonomních systémů**.

Většina skupinového provozu probíhá uvnitř PIM domény, ovšem je třeba umožnit jí i komunikaci s okolním světem. V IPv4 si proto **jednotlivé PIM domény vzájemně vyměňují informace o svých skupinových zdrojích.** To umožňuje, aby vznikaly skupiny s příjemci a zdroji roztroušenými v několika PIM doménách. Pro vzájemnou výměnu informací o existujících skupinách a zdrojích slouží ***Multicast Source Discovery Protocol (MSDP)***, který má úlohu podobnou BGP ve světě individuálního směrování.

IPv6 díky svým dlouhým adresám může zařadit adresu shromaždiště přímo do skupinové adresy. Žádná výměna informací o dění uvnitř PIM domén pak není potřeba, kdokoli v celém Internetu se přímo z adresy skupiny dozví, kde je její shromaždiště a kam má tedy poslat žádost o příjem dat.

Práce na adaptaci protokolu MSDP pro IPv6 byly proto zastaveny a nahrazeny konceptem vložených adres.

PIM Sparse Mode (PIM-SM)

Je **nejpoužívanějším a i nejsložitějším z celé skupiny.** V různých situacích jsou nároky na co nejefektivnější distribuci skupinových dat odlišné, proto v sobě PIM-SM kombinuje několik odlišných mechanismů.

Jeho definice je v RFC 4601: *Protocol Independent Multicast – Sparse Mode (PIM-SM): Protocol Specification*.

Základní myšlenkou PIM-SM je, že **skupinově adresovaná data se doručují jen tam, kde si o ně někdo řekl**.

Používá k tomu **typ zpráv *Připojení (PIM Join)*,** jejichž prostřednictvím ohlašuje zájem o odběr skupiny směrovač, kterému se ohlásil alespoň jeden její příjemce.

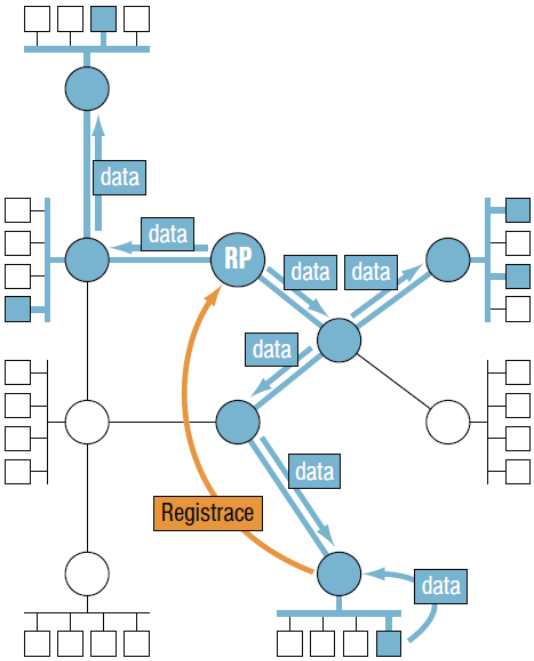
Kam žádosti o příjem skupiny adresovat? PIM-SM pro tento účel zavádí **speciální směrovač, označovaný jako *shromaždiště (rendezvous point, RP)***. Představuje místo v síti, kde si dávají dostaveníčko odesilatelé dat určité skupiny s jejich příjemci.

PIM-SM vytváří pro skupinu takzvaný *sdílený strom (shared tree)*, jehož kořenem je shromaždiště a větve dosahují do všech směrovačů, které se přihlásily k odběru skupiny.

Základní distribuce dat vypadá tak, že

- odesilatel skupinových dat pošle datagram
- jeho přilehlý směrovač jej zabalí do PIM zprávy nazývané z nevyzpytatelných důvodů *Registrace (Register)*a pošle na individuální adresu shromaždiště RP.
- na RP se příchozí datagram vybalí a rozešle sdíleným stromem příjemcům.

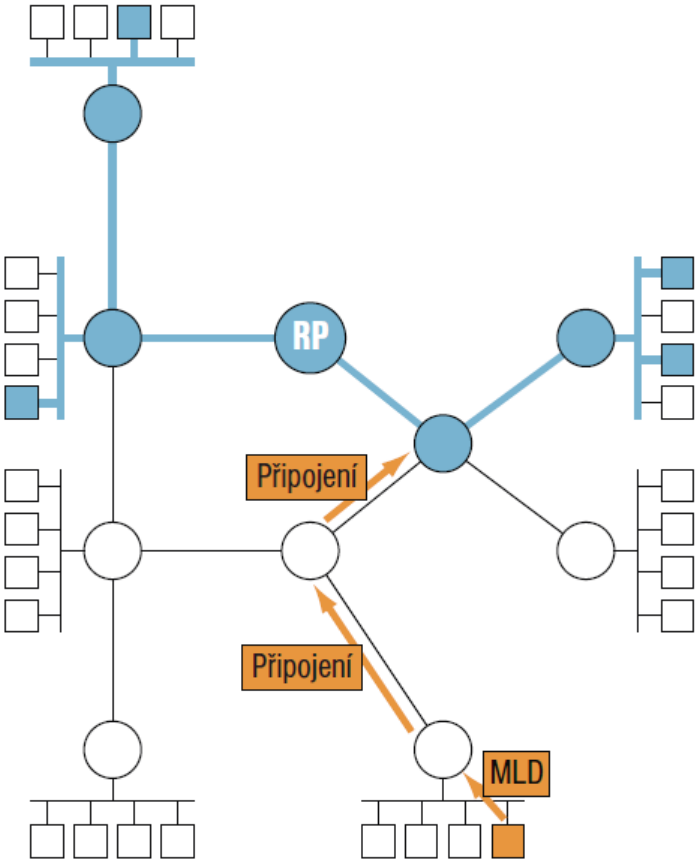
Sdílený strom je podle této představy jednosměrný a slouží k distribuci dat od RP k příjemcům, příklad je na obrázku



Když se směrovači protokolem MLD ozve zájemce o novou, zde dosud nepřijímanou skupinu, **směrovač pošle shromaždišti dané skupiny zprávu *Připojení*,** kterou žádá o zapojení do jejího sdíleného stromu. Skupina, do níž se hlásí, bývá označována jako (*,G). Tento zápis znamená, že odesilatel má zájem o pakety ze všech zdrojů (hvězdička) zaslané na skupinovou adresu G.

Připojení se posílá na individuální adresu RP. Směrovače po cestě si poznamenají, že do rozhraní, ze kterého přišla, mají do budoucna posílat danou skupinu. Jakmile zpráva dorazí ke směrovači, který již je součástí jejího sdíleného stromu, žádost o vstup se zahodí, protože právě byla naplněna. Celá posloupnost směrovačů, jimiž na své cestě prošla, se přidá do sdíleného stromu skupiny a bude se do budoucna účastnit předávání jejích dat.

Proces zapojení do sdíleného stromu



Dokud směrovač má přímé či nepřímé posluchače skupiny, posílá v určitých intervalech *Připojení*, aby u nadřízených občerstvil své členství ve sdíleném stromu. **Jestliže z určité větve dlouho nepřicházejí připojovací zprávy, odřízne se.**

Při vstupu do sdíleného stromu i rozesílání dat hrají důležitou roli směrovače poblíž příjemců a zdrojů dat. Počítače jsou dnes ale připojeny sítěmi s mnoha účastníky a do nich může být zapojeno i několik směrovačů. Je potřeba určit ten jediný. PIM-SM pro tento účel definuje mechanismus, kterým si skupina směrovačů připojených ke stejné lince vybere mezi sebou tak zvaný ***zodpovědný směrovač (designated router, DR)*, který zastupuje tuto síť**

Už z jednoduchého příkladu je patrné, že rozesílání dat sdíleným stromem je neefektivní. Do míst ležících poblíž zdroje se dostávají datagramy přes RP(takže z Olomouce do Ostravy se jezdí přes Prahu). **PIM proto obsahuje několik optimalizačních mechanismů.**

Konec registrace (Register-Stop)- vychází z toho, že odesilatel skupinových dat bývá zároveň příjemcem dané skupiny. Jedna větev sdíleného stromu proto vede směrem k němu a **data po ní proudí dvakrát To se dá odstranit**.

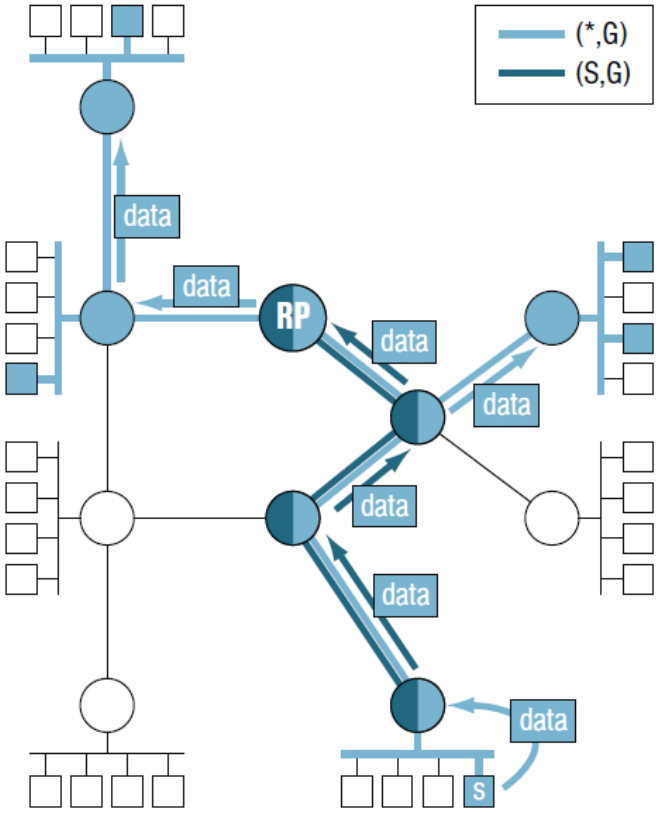
Shromaždiště po obdržení registrovaného datagramu ze zdroje S pro skupinu G pošle zodpovědnému směrovači, který odeslal *Registraci*, žádost o vstup do stromu (S,G).

Jedná se o nový distribuční strom pro datagramy směřující na skupinovou adresu G odeslané ze zdroje S. **Použije k tomu zprávu *Připojení*** a stejně jako ve výše popsaném případě **sdíleného**

stromu pro (*,G) se nyní vytvoří větev stromu pro (S,G). Směrovače na ní budou zpravidla také součástí sdíleného stromu (*,G) a budou datagramy přicházející stromem (S,G) předávat do stromu (*,G).

Jakmile shromaždišti dorazí první datagram ze stromu (S,G), **pošle zodpovědnému směrovači pro S zprávu *Konec registrace*,** kterou říká „už mi neposílej registrované datagramy, dostávám je přímo“.

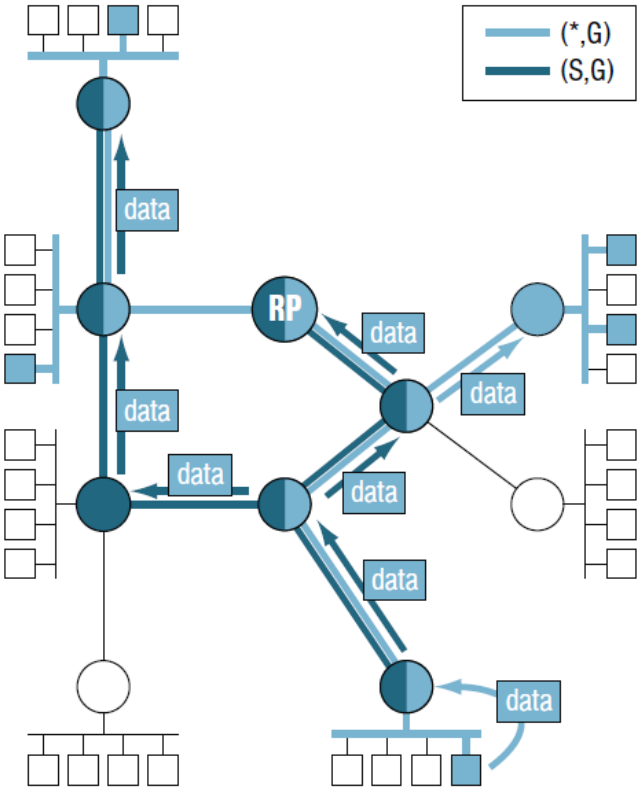
Distribuce dat po provedení Register-Stop



Efektivita šíření dat se tak výrazně zlepší, ale pro některé příjemce může být cesta přes RP slušnou oklikou. Směrovač zapojený do sdíleného stromu skupiny G se může rozhodnout5 zapojit se přímo do stromu (S,G), jehož kořenem je zodpovědný směrovač zdroje S.Pošle tedy PIM zprávu *Připojení* pro vstup do stromu (S,G) směrem ke zdroji S. Stejně jako dříve to znamená, že se někde (možná až v kořeni) napojí na tento strom - ***strom nejkratších cest(Shortest-Path Tree, SPT)***.

Po rozšíření začne některý směrovač v něm dostávat datagramy dvojmo – jednou přímo od zdroje stromem nejkratších cest, podruhé se zpožděním od RP sdíleným stromem.

Na to reaguje odříznutím ze sdíleného stromu pro tento konkrétní zdroj. Pošle svém **nadřazenému uzlu ve sdíleném stromě zprávu *Odříznutí (Prune)*** pro (S,G), kterou říká „data do skupiny G pocházející od S už mi neposílej“.



Strom nejkratších cest

Pokud nadřízený nemá jiného odběratele, předá zprávu dál nahoru – ze sdíleného stromu se tak pro zdroj S odřízne celá nepotřebná větev[6]. Situaci naší ukázkové skupiny poté co směrovač vlevo nahoře vstoupil do stromu nejkratších cest znázorňuje výše uvedený obrázek

PIM-SM je poměrně minimalistický ohledně počtu různých typů zpráv. Vystačí s pouhými šesti základními typy.
Složitá jsou pravidla pro jejich vysílání a zpracování – kdo komu a za jakých podmínek posílá kterou z nich a jak se má chovat její příjemce

Určení shromaždiště

Aby PIM-SM dobře fungoval, **musí umět najít adresu shromaždiště pro každou skupinu.** Navíc musí být tato adresa v rámci dosahu skupiny jednoznačná, jinak by vznikaly oddělené sdílené stromy.

PIM-SM nabízí několik řešení:

- **statická konfigurace RP pro jednotlivé skupiny,** kterou podle RFC 4601 musí podporovat každý směrovač implementující PIMSM.

- **vložení adresy RP přímo do skupinové adresy** (tzv. embedded RP). Na rozdíl od ostatních je tato **možnost dostupná jen pro IPv6,** protože v IPv4 adrese na něco takového prostě není dost místa. Má příjemné vlastnosti, protože je na jedné straně dost pružná –umožňuje vytvářet RP podle potřeby – a zároveň jednoduchá. Z adresy skupiny se snadno sestaví adresa jejího RP a může se komunikovat.

- třetí varianta je nejsložitější. Definuje **postup pro dynamické určování shromaždišť jednotlivých skupin.** Jeho definici najdete v RFC 5059: *BootstrapRouter (BSR) Mechanism for Protocol Independent Multicast (PIM)*. Spočíváv tom, že směrovače v jedné PIM doméně si mezi sebou vyberou jednoho„rozhodčího“, označovaného jako *bootstrap router (BSR)*. U něj

se pak ostatní směrovače ucházejí o roli RP. BSR vybere shromaždiště pro jednotlivé skupiny a tuto informaci pak rozšíří všem směrovačům v PIM doméně.

RP je pro skupinu unikátní a představuje tak zároveň její slabé místo. Pokud přestane pracovat, zastaví se činnost sdíleného stromu (data do něj posílá jen RP) a skupina bude mít vážné potíže.

Jednou z možností, jak tento problém obejít, je realizovat shromaždiště skupinou směrovačů se společnou výběrovou adresou (anycast RP). Ze stromu se pak vlastně stane les, příjemci i vysílající jsou napojeni vždy na nejbližšího zástupce RP, který pak datové pakety předává i dalším dílčím RP k odeslání do jejich stromů. Podrobněji toto uspořádání popisuje RFC 4610: *Anycast-RP Using Protocol Independent Multicast (PIM)*.

Další skupinové směrovací protokoly jsou nad rámec tohoto seznámení a blíže v knize p.Satrapy.