

Rozesílání a směrování vícesměrového vysílání

Multicast Switching and Routing

Bc. Václav Černý

Diplomová práce
2013

 Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky
akademický rok: 2012/2013

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Václav ČERNÝ**
Osobní číslo: **A10313**
Studijní program: **N3902 Inženýrská informatika**
Studijní obor: **Počítačové a komunikační systémy**
Forma studia: **prezenční**

Téma práce: **Rozesílání a směrování vícesměrového vysílání**

Zásady pro vypracování:

1. Podrobně se seznamte s principy vícesměrového vysílání.
2. Popište využití vícesměrového vysílání pro přenosy IPTV.
3. Provedte konfiguraci všech nutných součástí sítě.
4. Vyhodnoťte datový provoz sítě a porovnejte s unicasty.
5. Popište připravenost řešení na přechod na IPv6.

Rozsah diplomové práce:

Rozsah příloh:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

1. S. DEERING a B. CAIN. RFC 1112: Host Extensions for IP Multicasting. In: *ietf.org: Network Working Group [online]. Stanford University, 1989 [cit. 2013-01-20]. Dostupné z: <http://tools.ietf.org/html/rfc1112>*
2. H. HOLBROOK, ARASTRA, INC., B. CAIN a B. HABERMAN. RFC 4604: Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source-Specific Multicast. In: *ietf.org: Network Working Group [online]. The Johns Hopkins University Applied Physics Laboratory, 2006 [cit. 2013-01-20].*
3. COTTON, M., L. VEGODA, D. MEYER. RFC 5771: IANA Guidelines for IPv4 Multicast Address Assignments. In: *ietf.org: Network Working Group [online]. Internet Corporation for Assigned Names and Numbers, 2010 [cit. 2013-01-20]. Dostupné z: <http://tools.ietf.org/html/rfc5771>*
4. STEVENS, Richard. TCP/IP illustrated. Vyd. 1. Boston: Addison-Wesley, 1994, 576 s. ISBN 02-016-3346-9.
5. IP Multicast: Benefits of IP Multicast. Cisco.com [online]. 2012 [cit. 2013-01-20].

Vedoucí diplomové práce:

Ing. Jiří Korbel, Ph.D.

Ústav počítačových a komunikačních systémů

Datum zadání diplomové práce:

26. února 2013

Termín odevzdání diplomové práce:

31. května 2013

Ve Zlíně dne 26. února 2013


prof. Ing. Vladimír Vašek, CSc.
děkan




prof. Ing. Karel Vlček, CSc.
ředitel ústavu

ABSTRAKT

Ve své diplomové práci se zabývám popisem, rozbořem a aplikací vícesměrového vysílání datagramů skupině příjemců, které v prostředí telekomunikačních sítí vhodně doplňuje obvyklé zasílání datagramů jedinému příjemci.

Díky výraznému zvýšení efektivity přenosu dat po síti otevírá vícesměrové vysílání dveře k nasazení datově náročných síťových aplikací i tam, kde by to při použití jednosměrového vysílání nebylo z kapacitních důvodů nikdy možné. V tomto případě se jedná konkrétně o realizaci IPTV (Internet Protocol Television) v síti místního poskytovatele internetu.

V práci dále popisuji požadavky na síťová zařízení a jejich konfiguraci, včetně analýzy síťového provozu při použití vícesměrového vysílání a vyhodnocení efektivity přenosu dat ve srovnání s běžným jednosměrovým vysíláním.

V neposlední řadě se dotýkám i problematiky nasazení vícesměrového vysílání datagramů adresovaných protokolem IPv6, včetně simulace jejich směrování.

Klíčová slova:

Multicast, IGMP, UDP, PIM, IPTV, IPv6

ABSTRACT

In my diploma thesis I deal with the description, analysis and application of multicast packet transmission for a group of recipients that in an environment of telecommunication network complements the conventional sending datagrams to a single recipient.

Due to the significant increase in the efficiency of data transfer over the network, multicast opens the door to a deployment of data-intensive network applications at the point, where the utilization of unicast would have been impossible because of the bandwidth constraints. In this case, the implementation of the Internet service provider's Television over IP service.

The paper further describes the requirements for devices on the network and their configuration, including network traffic analysis using multicast addressing and evaluate the efficiency of data transfer compared to conventional one directional transmission.

Finally, I also cover the issue of deployment of multicast datagrams addressed as IPv6, as well as the simulation of routing.

Keywords:

Multicast, IGMP, UDP, PIM, IPTV, IPv6

Děkuji za podporu všem zúčastněným, zejména pak vedoucímu práce panu Ing. Jiřímu Korbelovi Ph.D. za odborné vedení, cenné rady a veškerou pomoc v průběhu tvorby této práce.

Děkuji také svému zaměstnavateli, který mi umožnil využít firemní síťovou infrastrukturu.

Prohlašuji, že

- beru na vědomí, že odevzdáním diplomové/bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomová/bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové/bakalářské práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji diplomovou/bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou/bakalářskou práci nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- beru na vědomí, že pokud bylo k vypracování diplomové/bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové/bakalářské práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové/bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně

.....

podpis diplomanta

OBSAH

ÚVOD.....	1
I. TEORETICKÁ ČÁST	3
1 MULTICAST	4
1.1 MAPOVÁNÍ MULTICASTOVÝCH IP ADRES NA MAC ADRESY.....	4
1.2 ADRESOVÁNÍ MULTICASTU	6
1.3 SMĚROVÁNÍ MULTICASTU	8
1.3.1 <i>Reverse Path Forwarding (RPF)</i>	8
1.3.2 <i>Režimy funkce směrovacích protokolů</i>	11
1.3.3 <i>Protocol Independent Multicast (PIM)</i>	12
1.3.4 <i>DVMRP, MOSPF a BGMP</i>	20
1.4 MULTICAST NA PŘÍSTUPOVÝCH ZAŘÍZENÍCH	21
1.4.1 <i>Internet Group Management Protocol (IGMP)</i>	21
1.5 APLIKACE A ROZŠÍŘENÍ	25
2 OSTATNÍ ZPŮSOBY VYSÍLÁNÍ.....	26
2.1 UNICAST.....	27
2.1.1 <i>Adresování</i>	27
2.1.2 <i>Směrování</i>	27
2.1.3 <i>TCP/UDP</i>	27
2.1.4 <i>TCP Three-Way Handshake</i>	27
2.2 BROADCAST.....	29
2.2.1 <i>ARP - Address Resolution Protocol</i>	30
2.2.2 <i>DHCP - Dynamic Host Configuration Protocol</i>	30
2.2.3 <i>Broadcast storm</i>	30
2.3 ANYCAST	31
2.4 GEOCAST	32

II. PRAKTICKÁ ČÁST	33
3 VYUŽITÍ MULTICASTU PRO PŘENOSY IPTV	34
3.1 MEDIA DELIVERY INDEX (MDI)	34
3.1.1 <i>Delay Factor (DF)</i>	34
3.1.2 <i>Media Loss Rate (MLR)</i>	35
3.2 KVALITATIVNÍ POŽADAVKY NA PŘENOS IPTV	36
3.2.1 <i>Propustnost</i>	36
3.2.2 <i>Latence</i>	37
3.2.3 <i>Ztrátovost</i>	37
3.2.4 <i>Změna pořadí paketů</i>	37
3.3 QUALITY OF SERVICE (QOS).....	38
3.3.1 <i>DiffServ Code Points (DSCP)</i>	38
3.3.2 <i>Class of Service (CoS)</i>	39
3.3.3 <i>Mapování DSCP na CoS</i>	39
3.4 MULTICASTOVÁ IPTV PLATFORMY.....	40
3.4.1 <i>Nangu.TV</i>	40
3.4.2 <i>Set top box</i>	40
3.4.3 <i>Nelineární služby</i>	40
3.4.4 <i>Šifrování přenosu</i>	40
3.5 SÍŤOVÁ TOPOLOGIE.....	41
3.5.1 <i>PIM-DM doména</i>	42
3.5.2 <i>IGMP doména</i>	44
3.5.3 <i>Sondy</i>	45
3.6 SÍŤOVÉ PRVKY A JEJICH KONFIGURACE	46
3.6.1 <i>PIM-DM doména</i>	46
3.6.2 <i>IGMP doména</i>	48
3.6.3 <i>Sondy</i>	50
3.7 ANALÝZA PROVOZU	51
3.7.1 <i>PIM-DM doména</i>	51
3.7.2 <i>IGMP doména</i>	56
4 MULTICAST VS. UNICAST	62

5	MULTICAST A IPV6	63
5.1	ADRESOVÁNÍ MULTICASTŮ V IPV6	63
5.1.1	<i>Mapování IPv6 adresy na MAC adresu.....</i>	66
5.1.2	<i>Multicast Listener Discovery (MLD).....</i>	66
5.2	SMĚROVÁNÍ MULTICASTŮ V IPV6.....	69
5.2.1	<i>Statická konfigurace RP.....</i>	69
5.2.2	<i>Mapovací agent BSR.....</i>	69
5.2.3	<i>Embedded RP.....</i>	70
5.3	SIMULACE SMĚROVÁNÍ IPV6 MULTICASTU V PROSTŘEDÍ CISCO IOS	71
5.3.1	<i>Platforma.....</i>	71
5.3.2	<i>Verze IOS a HW vybavení.....</i>	71
5.3.3	<i>Síťová topologie a nastavení prvků.....</i>	72
5.3.4	<i>Určení a nastavení Embedded RP</i>	73
5.3.5	<i>Analýza provozu.....</i>	74
5.4	MOŽNOSTI IMPLEMENTACE V SÍTI LOKÁLNÍHO ISP	77
5.4.1	<i>Stav alokace IPv4 a cesty řešení</i>	77
5.4.2	<i>Kompatibilita IPv6 multicastu se síťovými prvky.....</i>	78
5.4.3	<i>Aplikace a budoucnost.....</i>	80
	ZÁVĚR	81
	CONCLUSION	82
	SEZNAM POUŽITÉ LITERATURY	83
	SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK	86
	SEZNAM OBRÁZKŮ	87
	SEZNAM TABULEK.....	89
	SEZNAM PŘÍLOH	89

ÚVOD

V průběhu deseti let se z původně z čistě výzkumného nástroje stal internet každodenní součástí lidského života. Vytvořil nový věk, ve kterém jsou informace šířeny volně a všem bez rozdílu. Tato celosvětová síť změnila nejen způsob lidské komunikace, ale i to, jak pracujeme, bavíme se a myslíme. Donutila nás přehodnotit některé ze základních principů, které jsme po generace brali jako samozřejmost.

Téměř každý z možných aspektů lidské komunikace je svázaný s internetem. Převažuje samozřejmě komunikace prostřednictvím textu a statického obrazu, následuje zvuk a pohyblivý obraz. Právě video ale v posledních letech výrazně nabývá na významu. Jedná se totiž o nejpřirozenější formu komunikace, jakou si v prostředí internetu dnes umíme představit.

V porovnání s ostatními způsoby přenosu informace ale přenos videa také konzumuje řádově více přenosové kapacity. Zatímco problém propustnosti “poslední míle” byl již v minulosti vyřešen nástupem xDSL (varianty Digital Subscriber Line) technologie a rozvojem FTTx (Fiber To The Home / Fiber To The Building) v posledních letech, komplikace na straně poskytovatelů internetového připojení stále přetrvávají.

Z historických důvodů a vinou jen okrajového zastoupení ostatních druhů vysílání je k doručování obsahu ve vnitřních sítích poskytovatelů internetu využíván až na výjimky pouze unicast. Zatímco takové řešení funguje dobře pro přenos informace ve formě textu a statických obrazových dat, selhává v případě videa, jelikož škáluje lineárně se vzrůstajícím datovým provozem, který přenášení obrazových dat ve velkém generuje.

Přestože vícesměrové vysílání nabízí řešení tohoto problému, rychlost jeho nasazování je spíše pomalá. Důvodů je několik, společným jmenovatelem se ale zdají být zvýšené náklady spojené s nákupem vhodného technického vybavení, především pak ale absence služby, která by prostředky investované do budování nové infrastruktury přinesla nazpět. Již zmíněné multimediální služby bylo totiž donedávna možné provozovat i na stávající unicastové infrastruktuře. Zákazník tak v důsledku nepoznal rozdíl a nebyl tedy logicky ochoten hradit zvýšené náklady spojené s implementací vícesměrového vysílání.

Tato situace se ale začíná měnit. Jak se v posledních letech dramaticky zvyšují požadavky nejen na kvalitu přenosu, ale i na počet přijímaných kanálů, roste také datová náročnost přenášených video streamů. Ba dokonce roste do takové míry, že už nemůže být ve velkém měřítku přenášena přes stávající unicastová řešení. Vznikají tak specializované multicastové televizní platformy přinášející poskytovatelům možnost dodat zákazníkům kompletní balík televizních služeb včetně placených prémiových kanálů a nelineárních služeb, které poskytovatelům přinášejí kýžený zisk. Těmito prostředky pak poskytovatelé mohou zpětně financovat úpravu své infrastruktury potřebnou pro implementaci multicasu.

Kruh se tedy uzavírá. Zdá se, že po více než pětadvaceti letech po první standardizaci vícesměrového vysílání vzniká reálná aplikace, která nejenže dokáže využít škálovací schopnosti multicasu, ale také poskytovatelům vrací finance, které do jejího nasazení vložili.

Multicast má ale i svou odvrácenou stranu. Přes všechny své pozitivní vlastnosti přináší i vyšší nároky na správu a proškolení obsluhy. Pro úspěšné nasazení a dlouhodobé využívání vícesměrového vysílání na síti je nutné zbavit se některých základních axiomů, na kterých síťový svět stavěl poslední desetiletí. Při implementaci je nutné správně nakonfigurovat a udržovat většinu klíčových L3 prvků na síti, upravit firewall a mít dobrou kontrolu nad všemi předávacími uzly sítě tak, aby se multicast nešířil nekontrolovaně do dalších sítí či internetu vůbec.

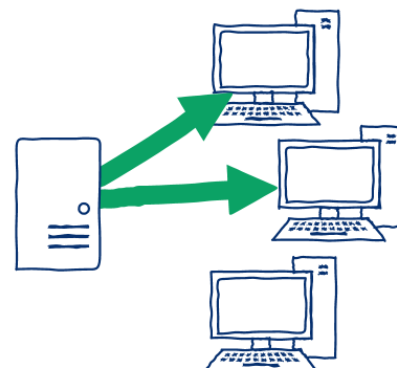
I tak ale vícesměrové vysílání v prostředí internetu nabízí jedinečnou možnost přivedení bohatého multimediálního obsahu z nahrávacích studií až do klientských domácností. A nejen to. Na rozdíl od klasického příjmu signálu přes satelitní parabolu dokáže multicast ve spojení s unicastem poskytnout i obousměrnou komunikaci, která je klíčová pro rozvoj nových aplikací interaktivní domácí zábavy dnes i v budoucnu.

I. TEORETICKÁ ČÁST

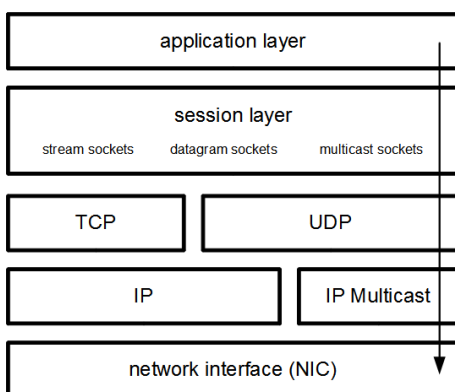
1 MULTICAST

Metoda vícesměrového vysílání byla poprvé definována už roku 1986 v RFC 1112. Doplnuje mezeru, která vznikla mezi jednosměrovým a všesměrovým vysíláním tak, že adresuje nikoliv koncové účastníky, ale celé skupiny příjemců, tedy multicastové skupiny.

Obecně se členem (či posluchačem) multicastové skupiny stává každý účastník, který vyjádří zájem přijímat pakety adresované této skupině. Zatímco členů skupiny může být $n-1$, zdroj bývá obvykle pouze tento jediný. Zdroj tedy nemůže být členem dané multicastové skupiny – poskytování a přijímání paketů z/do multicastové skupiny se vzájemně vylučuje. Zdroj navíc ani nezná konkrétní příjemce dat dané multicastové skupiny, jejímž je zdrojem.



obr. 1 - multicast



obr. 2 - multicast ISO/OSI

V případě IPv4 je multicast v podstatě rozšířením jeho třetí vrstvy podle ISO/OSI modelu. Nová verze protokolu IPv6 již multicast zahrnuje, a nejen to. Vícesměrové vysílání v nové verzi protokolu IP dokonce zcela nahrazuje všesměrové vysílání. Téma je dále diskutováno v páté kapitole *Multicast a IPv6*.

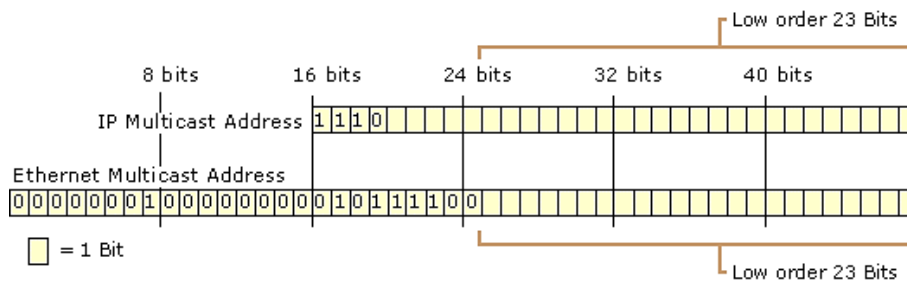
Vícesměrové vysílání je dále logicky spoléhá na UDP datagramy. Jedná se tedy o nezabezpečený a negarantovaný přenos dat.

1.1 Mapování multicastových IP adres na MAC adresy

Ethernetové rámce multicastových datagramů jsou na linkové vrstvě snadno rozeznatelné – nultý bit nultého oktetu (tedy prvního bajtu) je vždy nastaven na 1. Toto platí pouze s jedinou výjimkou, a to v případě broadcastového rámce, jehož cílová adresa je FF:FF:FF:FF:FF:FF - v binární podobě jsou tedy všechny bity všech šesti bajtů MAC adresy nastaveny na 1.

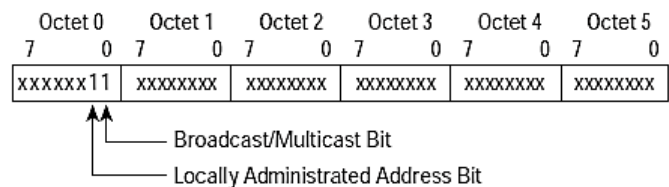
V případě IP unicastu neexistuje přímé mapování mezi IP adresou a MAC adresou – síťová adresa na druhé vrstvě ISO/OSI modelu je získávána dynamicky protokolem ARP (*Address Resolution Protocol*).

V případě IP multicastu bylo ale využito přímé mapování síťové IP na MAC adresu [1]. Všechny multicastové rámce tak začínají 25bitovým prefixem 01:00:5E. Zbýlých 23 bitů je přeneseno z IP adresy. Vzhledem k tomu, že IP rozsah multicastových adres má pevný 4bitový prefix 1110 (odpovídá původní třídě D), zbývá 5 bitů, které z IP do MAC adresy nejsou nijak mapovány. V důsledku tedy existuje 2^5 IP adres, které jsou namapovány na stejnou MAC adresu.



obr. 3 - statické mapování multicastových MAC adres [1]

Proto např. multicastové adrese z privátního rozsahu **239.1.1.1**, binárně **01:00:5e:01:01:01**, na druhé vrstvě ISO/OSI modelu odpovídá i **224.1.1.1**, **224.129.1.1** a dalším 29ti IP adresám.



obr. 4 – bitový zápis multicast MAC adresy [2]

Celkem tedy 32 (2^5) multicastových IP adres sdílí jedinou multicastovou MAC adresu.

Důvod je přitom přinejmenším úsměvný. Když Steve Deering na počátku devadesátých let pracoval na úkolu vytvoření mechanismu, který by umožnil vícesměrové vysílání [2], uvědomil si, že k realizaci přímého mapování MAC adres potřebuje pokrýt prostor 28bitů (IPv4 adresa 32 bitů – 4bit prefix 1110 definující multicast). Cena přidělení 24bit bloku od Internet Assigned Number Authority (IANA) přitom tehdy vycházela na 1000\$ a k pokrytí 28 bitů jich bylo zapotřebí celkem 16 (2^4 bloků). Investice 16 000\$ byla ale tehdy Stevu Deeringovi zamítnuta nadřizeným. Musel se spokojit pouze s jedním 24bit blokem. Ba co víc, ze zakoupeného bloku dostal Deering pro svůj projekt přidělenou pouze polovinu adres, tedy 23bit. A tak se i dnes musíme smířit s nejednoznačným mapováním multicastových IP adres na MAC

1.2 Adresování multicastu

Na rozdíl od unicastové adresy, která popisuje jednoho unikátního klienta na síti, multicastová adresa definuje skupinu klientů, kteří se do této skupiny připojili. Pro adresování vícesměrového vysílání byl vyčleněn adresní IPv4 rozsah od **224.0.0.0** do **239.255.255.255** (všechny multicastové adresy začínají bity *1110*, odpovídají tedy třídě D dnes již nepoužívaného systému adresování podle tříd) [3]. Stejně jako u rozsahu přiděleného pro adresování jednosměrového vysílání, i výše uvedený rozsah se sestává z několika bloků rezervovaných pro speciální použití:

- 224.0.0.0/24 rozsah rezervovaných lokálních multicastových adres
- globální adresy:
 - 224.0.1.0/24 rozsah internetové správy (Internet Control Block)
 - 224.2.0.0/16 Session Announcement/Description Protocol (SAP, SDP)
 - 232.0.0.0/8 Source-Specific Multicast (SSM) (RFC 4607)
 - 233.0.0.0/8 staticky přidělené GLOP adresy (RFC 3180)
 - 234.0.0.0/8 adresy přiřazené k veřejným unicast rozsahům (RFC 6034)
- 239.0.0.0/8 adresy s omezeným dosahem (RFC 2365)

Rezervované adresy

Rezervované lokální adresy jsou využívány převážně multicastovými směrovacími protokoly. Pakety jsou omezeny hodnotou Time To Live (TTL) pouze na lokální segment, tedy 1. Konkrétně jsou to např.:

- 224.0.0.1 všichni účastníci na podsíti
- 224.0.0.2 všechny směrovače na podsíti
- 224.0.0.5(6) OSPF směrovače (designované - DRs)
- 224.0.0.12 DHCP server/relay agenti
- 224.0.0.13 Protocol Independent Multicast (PIM) (verze 2)
- 224.0.0.18 Virtual Router Redundancy Protocol (VRRP)
- 224.0.0.22 IGMP

Globální adresy

V globálním měřítku (*Globally Scoped Addresses*) jsou pak multicastové adresy přidělovány v rozsahu **224.0.1.0** až **238.255.255.255**.

- *Internetwork Control Block* adresy jsou rezervovány pro využití multicastovými aplikacemi pod správou IANA, např. multicastová adresa 224.0.1.1, která je vyhrazena pro synchronizaci času přes Network Time Protocol (NTP).
- *Session Announcement Protocol* (SAP) a *Session Description Protocol* (SDP) jsou experimentální protokoly informující o existenci dalších multicastových relací. Využití nalézají mimo jiné také v IPTV aplikacích, viz. třetí kapitola – Využití multicasu pro přenosy IPTV.
- Další /8 blok je vyčleněn pro využití Source-Specific Multicast (SSM), metodu zvyšující bezpečnost přenosu dat omezením počtu zdrojů na specifickou IP adresu odesílatele.
- Ke každému autonomnímu systému o rozsahu /16 je automaticky přiděleno /24 statických GLOP adres, které mohou být vlastníkem AS volně využity. 16bit číslo AS je přitom dosazeno do prostředních dvou oktětů rozsahu 233.0.0.0/8 ve tvaru 233.x.x.0 až 233.x.x.255.
- Podobně jako u GLOP adres, další /8 rozsah (konkrétně 234.0.0.0/8) je opět přidělen k existujícím unicastovým rozsahům, tentokrát jako jedna multicastová adresa pro každý /24 rozsah. Toto řešení škáluje lépe než GLOP adresy.

Adresy s omezeným dosahem

Adresy s omezeným dosahem (*Administratively Scoped Addresses*) jsou ekvivalentem privátních unicastových rozsahů s vyčleněným IP rozsahem 239.0.0.0/8. Mohou být tedy využívány opakovaně v různých sítích, zároveň je ale nutné zajistit, aby se takto adresované datagramy nešířily mimo AS poskytovatele nebo jinou omezenou doménu.

Hraniční PIM-DM router vždy odmítá záplavu z těchto adres (*PIM Prune*), stejně tak hraniční PIM-SM router odmítá připojování (*PIM Join*) na tyto adresy.

Z IP rozsahu 239.0.0.0/8 je dále vyčleněn rozsah **239.255.0.0/16**, který představuje **lokální adresy s omezeným dosahem** (*The IPv4 Local Scope*, RFC 2365) a **239.192.0.0/14**, který představuje adresy s omezeným dosahem vyčleněné pro použití v organizacích (*The IPv4 Organization Local Scope*, RFC 2365).

1.3 Směrování multicastu

Hlavní funkcí multicastových směrovacích protokolů je zajištění správného nastavení portu routeru k odesílání multicastových datagramů (tzv. *Multicast Forwarding State*), dále také výměna informací o stavu jejich směrování s dalšími routery.

Protože může existovat (a zpravidla také existuje) více příjemců multicastového streamu, tedy **posluchačů** (*receiver/listener*) jedné multicastové skupiny, cesta multicastových paketů se obvykle dělí do tzv. **větví** (*branches*). Systém těchto větví je pak **distribuční strom** (*distribution tree*). Data proudící ke zdroji multicastových dat, tedy **kořenu** distribučního stromu (*distribution tree root*) jsou označována jako **upstream**, data proudící směrem k posluchači pak jako **downstream** [4].

Routery přitom udržují *multicast forwarding state* pro každou multicastovou skupinu na každém logickém rozhraní daného routeru. Příchozí rozhraní bývá často označováno jako **IIF**, zatímco seznam odchozích rozhraní pro multicastovou skupinu bývá označován jako **OIL**. OIL přitom nabývá hodnot 0 až N, kde N značí celkový počet logických síťových rozhraní routeru.

Multicast forwarding state je v routerech značen většinou jako **(S,G)** nebo **(*G)**. Zatímco „**S**“ (*source*) odkazuje na unicastovou adresu zdroje, „**G**“ reprezentuje danou multicastovou skupinu. V IP hlavičce multicastového paketu je tedy pole „**S**“ zastoupeno v zdrojové IP adrese a pole „**G**“ v cílové IP adrese.

Hvězdička v notaci **(*G)** je tzv. „expanzním znakem“ (*wild card*), který zachycuje situaci, kdy jakýkoliv zdroj zasílá data multicastové skupině **G**. Každá multicastová skupina totiž může mít i více zdrojů než jeden [4].

1.3.1 Reverse Path Forwarding (RPF)

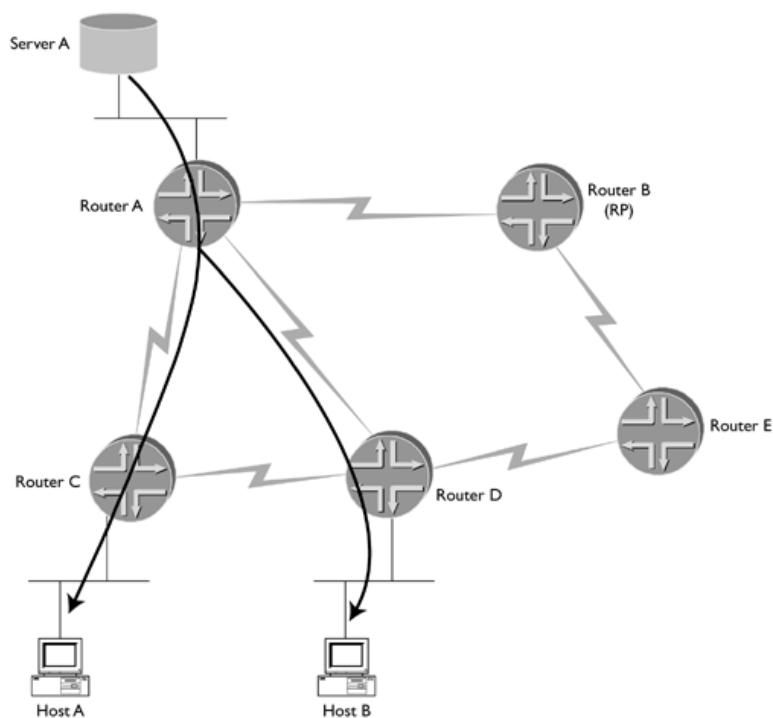
Reverse Path Forwarding představuje obecný koncept užívaný multicastovými směrovacími protokoly k zajištění funkční topologie bez smyček.

Směrování unicastových datagramů je založeno na porovnávání cílové IP adresy paketu s směrovací tabulkou daného routeru. V závislosti na záznamech v této směrovací tabulce router příchozí paket směruje na jedno ze svých síťových rozhraní. Unicastové pakety jsou tedy routovány od zdroje k cíli.

V případě multicastového směrování ale routery nastavují *multicast forwarding state* svých síťových rozhraní v opačném pořadí než v případě unicastu. Routery provádějí tzv. *Reverse Path Forwarding* (RPF) k určení tzv. *RPF interface*, tedy síťového rozhraní, které je topologicky nejbližší kořenu distribučního stromu [4]. *RPF interface* je zároveň *upstream* pro danou multicastovou skupinu.

V případě **Shortest Path Tree (SPT)** je kořen distribučního stromu zároveň i jeho zdrojem. Pokud tedy koncový router obdrží od koncového klienta požadavek k odběru dat z určité multicastové skupiny, router provede tzv. *RPF check*, porovná tedy svoji směrovací tabulku se zdrojovou IP multicastové skupiny, ke které se chce klient připojit. Výsledkem je určení síťového rozhraní, které je topologicky nejbližší multicastovému zdroji, tedy *RPF interface*.

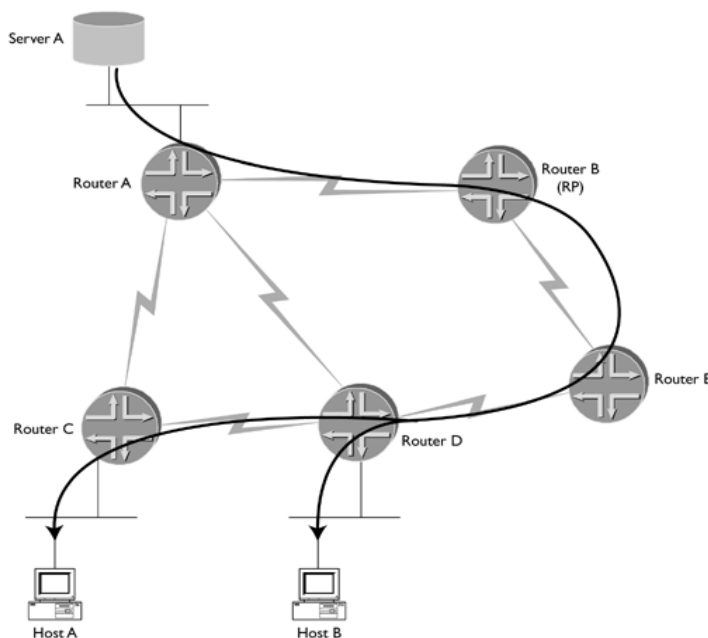
Na *RPF interface* router odešle tzv. *(S,G) join message*, tedy zprávu, kterou upstream routeru oznamuje svůj zájem o připojení do dané multicastové skupiny daného zdroje. Upstream router, který tuto zprávu obdrží jednak přidá příchozí rozhraní do OIL seznamu, dále také provede další *RPF check* pro určení *RPF interface*. Stejný postup opakuje každý upstream router až do chvíle, dokud nenarazí na router, který je přímo připojen k multicastovému zdroji nebo router, který má již multicast forwarding state pro danou multicastovou skupinu aktivní.



obr. 5 - Shortest Path Trees [4]

Vznikne tak nová větev v distribučním stromu s nejkratší cestou k multicastovému zdroji - Shortest Path Tree.

V případě sdíleného stromu není kořenem distribučního stromu samotný multicastový zdroj, ale jeden ze skupiny multicastových routerů na síti. Tento router bývá označován jako tzv. **rendezvous point (RP)**. Postup vytváření nové větve sdíleného stromu je velmi podobný jako v případě *Shortest Path Tree*, avšak s jediným rozdílem. *Join message*, která odchází přes *RPF interface* má formát $(*,G)$, jelikož router, který *RPF check* provedl zná ve skutečnosti pouze IP multicastové skupiny, nikoliv skutečnou IP multicastového zdroje. Dochází k opakovanému zasílání *Join messages* mezi routery po cestě směrem k RP podobně jako v případě SPT. Sdílený strom bývá také označován jako **Rendezvous Point Tree (RPT)** [4].



obr. 6 – Shared Distribution Trees [4]

K tomu, aby mohl RP správně směrovat multicastové datagramy do vlastní *rendezvous point tree*, musí je samozřejmě sám také přijímat. Proto sám provádí *RPF check* na potřebné multicastové skupiny, tentokrát však už mají *Join messages* tvar (S,G) , jelikož adresují přímo multicastový zdroj. Veškeré rozhodování se přitom děje na základě údajů z unicastové směrovací tabulky, jelikož jak adresa multicastového zdroje, tak i adresa RP jsou adresovány jako unicast.

Rendezvous point je praxi široce využíváno, zejména kvůli možnosti snazší administrace při změnách adres multicastových skupin nebo zařazení dalších multicastových zdrojů. Routery v *rendezvous point tree* jsou od těchto změn odstíněny. Na druhou stranu přináší taky možnost volby ne zcela optimální cesty, na rozdíl od *Shortest Path Trees*.

1.3.2 Režimy funkce směrovacích protokolů

Podle předpokladu četnosti a rozložení odběratelů datagramů na síti dělíme směrovací multicastové protokoly na tzv. **Dense** („husté“) a **Sparse** („řídke“) [4].

Zatímco *Dense* protokoly provoz implicitně směřují do všech segmentů sítě a až následně jej “prořezávají” (*pruning*) v případě, že v dané multicastové skupině není přítomen žádný příjemce, *Sparse* protokoly implicitně multicastové datagramy nesměrují nikam. Čekají až na přihlášení prvního účastníka do multicastové skupiny, tedy připojení další větve do vytvořeného sdíleného distribučního stromu, nebo vytvoření jeho samotného.

Dense mode

- Protocol Independent Multicast Dense Mode (PIM-DM) (RFC 3973)
- Distance Vector Multicast Routing Protocol (DVMRP) (RFC 1075)
- Multicast Open Shortest Path First (MOSPF) (RFC 1584)

Sparse mode

- Protocol Independent Multicast Sparse Mode (PIM-SM) (RFC 4601)
- Bidirectional PIM (BIDIR-PIM) (RFC 5015)
- Source-Specific Multicast (SSM) (RFC 3569)
- CBT (RFC2201) (RFC2189)

Další

- Border Gateway Multicast Protocol (BGMP) (RFC 3913)

1.3.3 Protocol Independent Multicast (PIM)

Protocol Independent Multicast označuje rodinu multicastových směrovacích protokolů, které v prostředí počítačových sítí realizují adresování a doručení jediného síťovém datagramu skupině příjemců, tedy vícesměrové vysílání. Na rozdíl od svých předchůdců je skutečně *protocol-independent*, jelikož při popisu topologie sítě nespolehá na vlastní techniku objevování sousedů, ale využívá záznamů z již existujících směrovacích tabulek. Tyto jsou nejčastěji plněny unicastovými směrovacími protokoly jako *Routing Information Protocol* (RIP), *Open Shortest Path First* (OSPF) nebo *Border Gateway Protocol* (BGP). V posledních letech se PIM (konkrétně PIM-SM) stal jasně dominantním protokolem pro směrování vícesměrového vysílání.

Definice pojmů a obecný přehled stavů protokolu PIM [5]

- **Rendezvous Point (RP)** – tzv. “místo setkání” pro odesílatele (zdroje) multicastových dat, tak i jejich příjemce. Představuje kořen pro sdílený distribuční strom (Shared Distribution Tree), který je budován směrem k posluchačům skupin. V první fázi PIM-SM rozbaluje unicastové pakety a následná multicastová data přeposílá příjemcům přes vybudovaný sdílený strom.
- **Designated Router (DR)** – tzv. „určený router”, který leží na hranici PIM domény a slouží jako prostředník mezi zprávami IGMP/MLD a PIM. DR v první fázi PIM-SM zapouzdřuje multicastové pakety do unicastových a přeposílá je na adresu RP. Pokud má více routerů síťové rozhraní z jediné podsíti, jeden z nich musí být zvolen jako DR (podle nejvyšší IP, která je součástí periodicky odesílaných PIM HELLO paketů).
- **Bootstrap Router (BSR)** – má privilegium zvolit jednoho z PIM routerů na doméně jako RP. V případě nedostupnosti původně zvoleného RP zvolí jeho náhradníka. Ostatní PIM routery informuje skrze tzv. *Bootstrap Messages*.
- **Register Messages** – multicastové datagramy, které jsou routovány od DR zdroje až k RP zapouzdřené jako unicast (směřované na unicast IP Rendezvous Point).
- **Assert Messages** – zprávy pro zajištění topologie bez kruhů. Jsou odesílány v případě, že router obdrží datagram adresovaný skupině G na stejném síťovém rozhraní, kterým datagramy skupině G odesílá.

V závislosti na konkrétní implementaci jsou všechny stavy protokolu PIM na routeru popsány v jeho tzv. *Tree Information Base* (TIB) [5], která představuje abstraktní definici popisující možné stavy protokolu PIM. V návaznosti na TIB se dále plní i *multicast forwarding table*.

Tree Information Base rozdělujeme na čtyři různé stavy [5]:

- **(*,***,RP**)** stav, ve kterém je udržován Rendezvous Point Tree (RPT) pro jakýkoliv zdroj všech skupin daného RP
- **(***,G**)** stav, ve kterém je udržován Rendezvous Point Tree (RPT) pro jakýkoliv zdroj konkrétní skupiny (G)
- **(S,G)** stav, ve kterém je udržován *source-specific tree* (SPT), tedy distribuční strom pro konkrétní zdroj (S) a konkrétní skupinu (G)
- **(S,G,rpt)** stav k udržení *source-specific* informace o zdroji (S), rendezvous point tree (RPT) a skupiny (G), typicky Prune zpáva při přechodu z RPT na SPT

Podle způsobu šíření a konstrukce distribučních stromů rozlišujeme několik funkčních módů.

PIM - Dense Mode

Tzv. “hustý mód”, specifikovaný v RFC 3973 [6]. Předpokládá častý výskyt příjemců napříč multicastovou doménou. Po výskytu nového multicastového zdroje skupiny *G Designated router* zdroj zahltí (*flood*) všechny své sousedy (*PIM neighbors*) a následně čeká na aktivní odmítnutí účastníků (*Prune*). Platnost tohoto odmítnutí vyprší standardně po třech minutách a záplava se opakuje. Pokud upstream router obdrží zprávu o obnovení streamu (*Graft*), nečeká na vypršení Prune časovače a změní ihned stav odesílání streamu z *Pruned* na *Forwarded*. Pokud naopak od svého souseda obdrží *State-Refresh* zprávu, prodlouží platnost Prune časovače o další 3 minuty – není tak třeba opět zaplavovat souseda všemi multicastovými streamy.

PIM-DM se vyznačuje nejsnazší implementací (absence *Rendezvous point*), ale také i nejhorším škálováním a výkonnostními parametry. V současné době je využíván pouze výjimečně, nejčastěji jako iniciační část PIM Sparse-Dense módu, tedy mechanismu automatického nalezení **auto-rp** (*auto rendezvous-point*), popř. na speciální aplikace, mezi které se počítá i IPTV.

Konkrétní popis proměnných a popisu chování protokolu PIM v režimu Dense mode následuje v druhé kapitole, *Využití multicasu pro přenosy IPTV*.

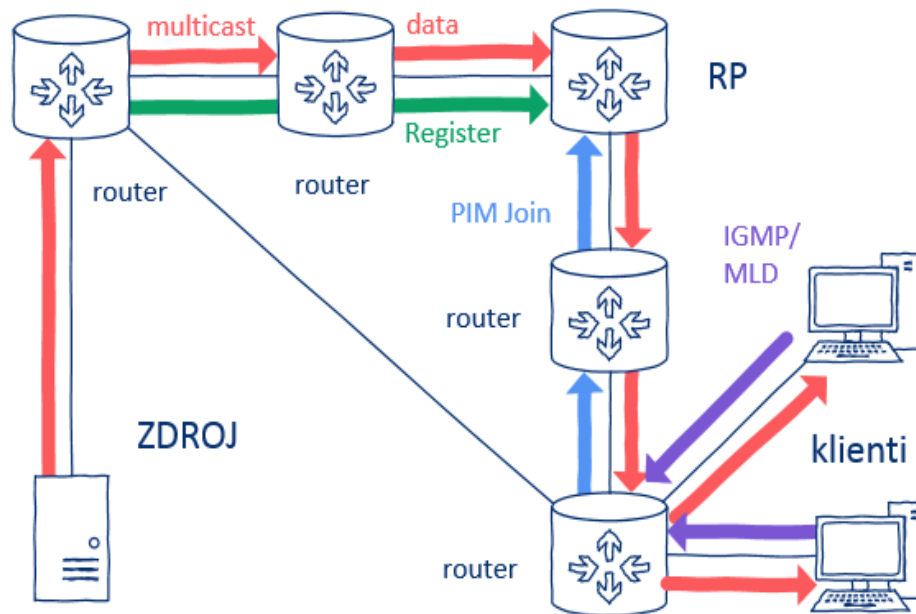
PIM - Sparse Mode

Tzv. “řídký mód”, specifikovaný v revidovaném RFC 4601 [5]. Předpokládá naopak jen ojedinělý výskyt posluchačů v multicastové doméně. Po výskytu multicastového zdroje nedochází k automatickému budování distribučního stromu, místo toho se čeká na *Join* zprávu od klienta, který tak musí nejdříve vyjádřit svůj zájem připojit se k multicastové skupině, tzv. **explicit join** model. Pro funkci je nutné definovat tzv. *Rendezvous-point* (RP), který se stává prostředníkem mezi multicastovým zdrojem a jeho odběrateli.

PIM-SM je nejčastěji používaný multicastový směrovací protokol, proto se také stal internetovým standardem. Ukázalo se totiž, že v praxi nejlépe výkonnostně škáluje, routery v *rendezvous point tree* navíc nemusí uchovávat informace (S,G) o každém multicastovém zdroji, pouze jednu (*,G) ukazující na *Rendezvous-point* (RP), což velmi zjednodušuje konfiguraci. Existují celkem tři vývojové fáze PIM-SM.

1. *RP tree*

Jak už bylo řečeno v jedné z předchozích kapitol, multicastové distribuční stromy jsou budovány od příjemce směrem ke zdroji dat. Prvním krokem v navázání komunikace je tedy vyjádření zájmu příjemce odebírat data z určité multicastové adresy, většinou prostřednictvím IGMP Report zprávy. Ta je zachycena na *Designated router* (DR) podsíti, ve kterém má příjemce své síťové rozhraní. DR tento požadavek přeposílá dále na adresu svého RP jako zprávu PIM Join. Tato zpráva odpovídá notaci (*,G), jelikož DR přeposílá požadavek o připojení se do sdíleného distribučního stromu určité multicastové skupiny (G), ale jakéhokoliv zdroje (*). Každý router, přes který tato PIM Join zpráva projde zároveň aktualizuje *multicast tree state* na svých rozhraních. Dříve nebo později tento PIM Join požadavek dorazí buď přímo k RP nebo na router, který již má vytvořen (*,G) záznam požadované skupiny – vytvoří se tak další větev budovaného distribučního stromu, tzv. *RP Tree* (RPT). Tento strom je sdílený – všechny zdroje využívají jeden a tentýž strom s kořenem v RP. Větve distribučního stromu jsou udržovány dokud jsou přeposílány periodické PIM Join zprávy. Větev zanikne v případě, že router nebo přímo RP obdrží (*,G) PIM Prune zprávu indikující, že v dané větvi se už nenachází žádný odběratel dané (G) multicastové skupiny. Další možností je vypršení časového limitu k obnově členství ve skupině – DR od účastníka neobdrží zprávu IGMP Report.



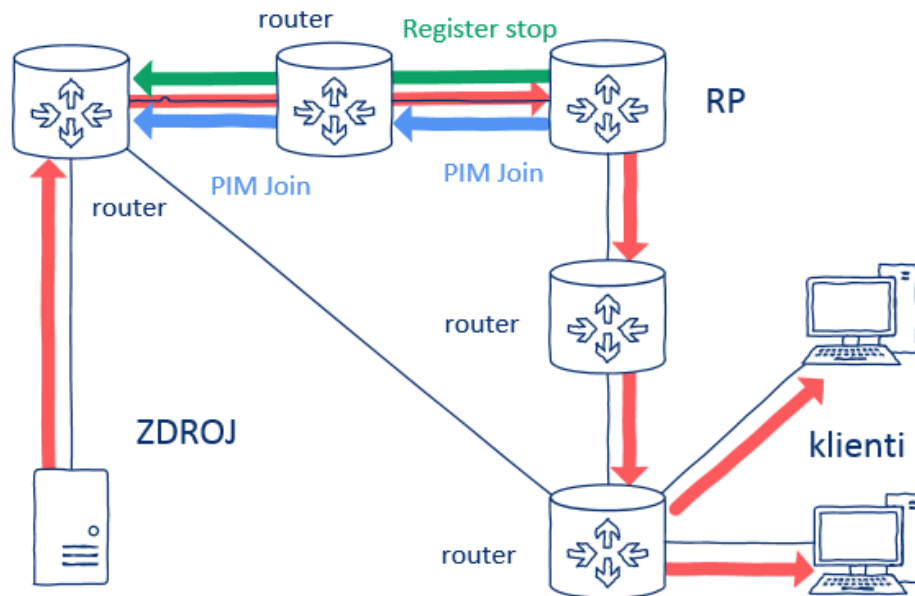
obr. 7 - první fáze PIM-SM – RP Tree

Multicastový zdroj přitom odesílá data určená pro danou multicastovou skupinu G okamžitě. DR tohoto zdroje tyto multicastové pakety zapouzdří do unicastových paketů a přešle je nejkratší cestou na adresu RP. Ten tyto pakety opět rozbálí a dále je distribuuje po sdíleném (*,G) RPT posluchačům skupiny G. Proces zapouzdření multicastových datagramů je znám jako *Registering* a zapouzdřené pakety jako *PIM Register* pakety.

2. Register-Stop

Koncept RP Tree ale nemusí být vždy zcela ideální. Nejenže je pouzdření multicastu na unicast poměrně výpočetně náročná operace, dále také může dojít k neoptimálnímu směrování datagramů. V první fázi PIM-SM totiž musí být všechny příchozí datagramy na RP přepouzdřeny zpět z unicastu (cílová adresa RP) na multicast (cílová adresa multicastové skupiny). Cesta na RP ale může datagramy multicastového zdroje zavést daleko od optimální cesty k příjemci.

Po obdržení zapouzdřeného PIM Register paketu od zdroje S, skupiny G, RP iniciuje zpětnou PIM Join zprávu pro specifický zdroj S. Jak tato zpráva postupuje zpět k DR multicastového zdroje, buduje postupně (S,G) distribuční strom (Shortest Path Tree) na routerech po cestě, až narazí buď přímo na DR s přímo připojeným zdrojem nebo na router, který již má daný (S,G) záznam vytvořen.



obr. 8 - druhá fáze PIM-SM – Register-Stop

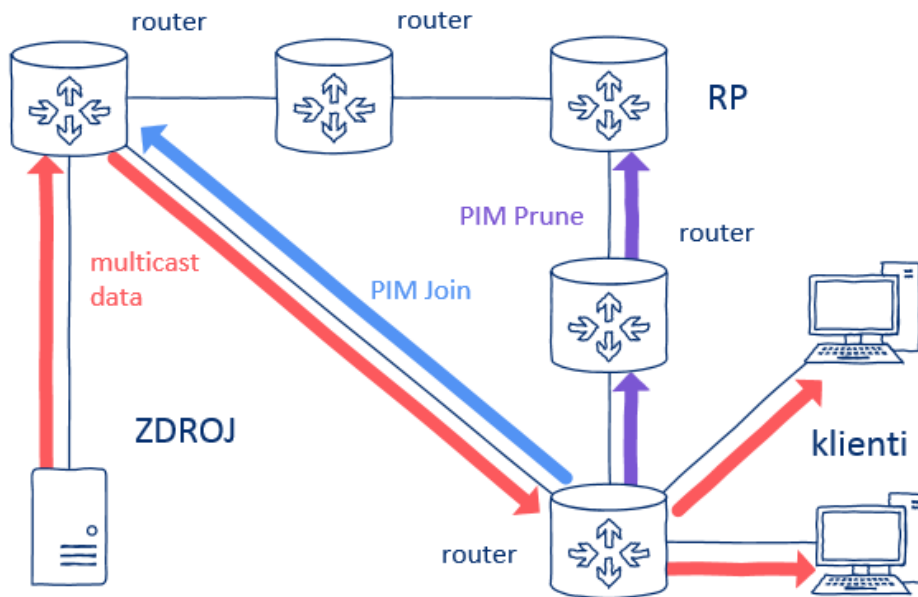
V tuto chvíli ale RP získává každý datagram zdroje dvakrát. Jednou zapouzdřený jako unicast a podruhé nativně přes nově vybudovaný SPT. Proto RP začne přijaté unicastové pakety aktivně zahazovat a DR multicastového zdroje zašle *Register-Stop* zprávu, po jejímž přijetí DR přestane multicastové datagramy zapouzdřovat a zasílat na adresu RP.

Druhá fáze tedy řeší problém pouzdření datagramů vybudováním (S,G) distribučního stromu, jenž má kořen v RP a větví se zpět ke zdrojům multicastových dat, tzv. Shortest Path Trees (SPT).

3. Shortest-Path Tree

Cesta datagramů přes fixně určené RP nemusí být vždy optimální. Z toho důvodu může router obsluhující příjemce (typicky DR) iniciovat PIM Join k nalezení nikoliv sdíleného distribučního stromu (*,G), ale nejkratšího možné cesty ke zdroji, tedy využití Shortest-Path Tree (SPT). Taková akce ale předpokládá znalost IP adresy, kterou standardně disponuje pouze RP.

V případě, že na DR příjemce dorazí dva stejné datagramy, jeden přes sdílený strom RPT, druhé přes nejkratší strom SPT, router datagram ze sdíleného stromu zahodí a navíc odešle zprávu PIM Prune směrem k RP - tzv. *(S,G,rpt) Prune* tak, aby došlo k odhlášení odběru dat ze sdíleného (*,G) stromu.



obr. 9 - třetí fáze PIM-SM – *Shortest-Path Tree*

PIM – Sparse-Dense Mode

Sparse-Dense mód je *Cisco* proprietární mód, který byl původně navržen tak, aby umožnil spojit to nejlepší z obou světů, tedy kombinovat *dense/sparse* přístup pro různé multicastové skupiny na jediném síťovém rozhraní. Často se používal ve spojení s **auto-rp** (*auto rendezvous-point*) – do doby automatického ustanovení *rendezvous-point* je multicast směřován v *Dense* módu, následně pak už v *Sparse* módu.

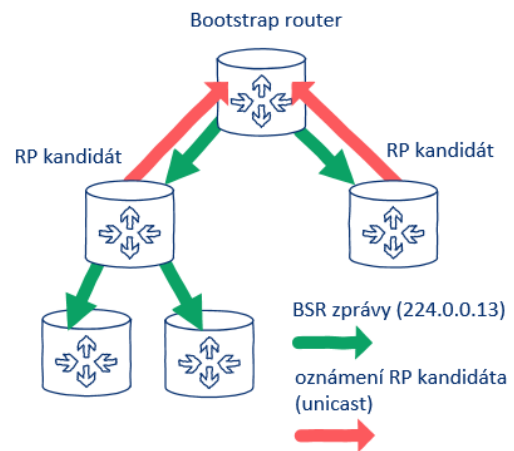
Auto-rp – opět původně *Cisco* proprietární funkce, kdy router automaticky vysílá do sítě oznámení o tom, že právě on by se rád stal tzv. *rendezvous-point*. Takových routerů ale může být na síti i více (RP pro konkrétní multicastové skupiny, apod.), proto je dále nutné zvolit i tzv. *mapping agent*

Mapping Agent (MA) - router určující, který z kandidátů na *rendezvous-point* se jím opravdu stane a pro kterou multicastovou skupinu. *Mapping Agent* může být stejný router jako *rendezvous-point*, ale z důvodu vyšší spolehlivosti je dobrým zvykem agenta delegovat fyzicky na jiný router.

Bootstrap router (BSR)

Kromě statického určení *rendezvous-point* v případě PIM-SM a využití *auto-rp* v případě módu PIM-SDM, je další možností využití protokolu **Bootstrap router (BSR)** [7].

BSR je otevřenou alternativou k *auto-rp*, která je využívána multicastovým směrovacím protokolem PIM-SM druhé verze. Samotná funkce je velmi podobná jako v případě *auto-rp*, pouze terminologie je odlišná. *Mapping Agent* je v případě BSR přímo *Bootstrap Router*.



obr. 10 – RP election process

Hlavním vylepšením oproti proprietárnímu *auto-rp* je nejen odbourání závislosti na platformě (nevyužívá se *cisco-rp-announce* (224.0.1.39) and *cisco-rp-discovery* (224.0.1.40) jako v případě *auto-rp*), především ale i počátečního použití PIM v módu *Dense* do ustanovení RP. Definice *rendezvous-point* IP je totiž součástí *BSR messages*, které jsou zase součástí *PIM messages*, kterými mezi sebou komunikují PIM routery ve společné lokální multicastové skupině 224.0.0.13.

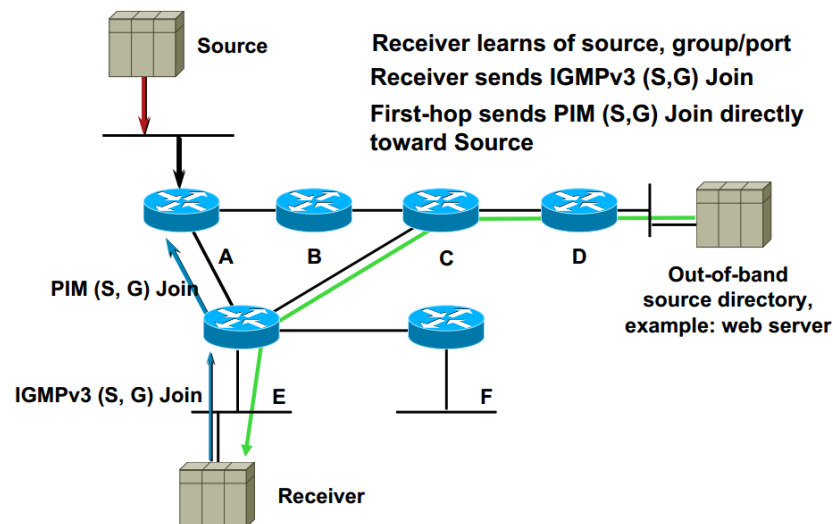
Bidirectional PIM (BIDIR-PIM)

Definováno v RFC 5015 [8]. *Bidirectional PIM* přináší možnost vytvářet *Bidirectional Shared Trees*, které mají kořen umístěný na IP adrese *Rendezvous-point*. Dochází tedy k eliminaci konkrétních (S,G) vazeb. Tento obousměrný provoz mezi multicastovým zdrojem a RP proto nevyužívá *Shortest Path Trees (SPTs)*, jako je tomu obvyklé u PIM-SM, a proto lépe škáluje v případě vysokého počtu zdrojů na multicastové doméně. Obousměrové (*,G) skupinové stromy jsou tedy využity jak pro přenos dat směrem od multicastového zdroje k RP, tak i od RP k příjemci.

BIDIR-PIM routery neprovádí *RFP check*. Na rozdíl od PIM-SM routerů proto musí být tyto routery schopny přijmout multicastový provoz i na více příchozích rozhraní. V praxi je jeho využití velmi omezené.

Source-Specific Multicast (PIM-SSM)

Definováno v RFC 3569 a RFC 4607 [9]. Běžný multicast, někdy také popisován jako *Any Source Multicast* (ASM), je založen na modelu komunikace *many-to-many*. Zdrojem dat v tomto modelu tedy může být kdokoli, známa je pouze IP adresa multicastové skupiny. Příjemce pouze signalizuje zájem přijímat data určité multicastové skupiny, ale o nalezení samotného multicastového zdroje už se sám nestará. Koncept *Source-Specific Multicast* přináší možnost odebírat data pouze z příjemcem explicitně specifikovaných multicastových zdrojů, nejčastěji v podobě *one-to-many*. Využívá se přitom pole INCLUDE IGMP Report zprávy verze 3.



obr. 11 - PIM-SSM [10]

Multicastový směrovací protokol, který podporuje implementaci SSM je odvozený od protokolu PIM, operujícím ve sparse módu. Proto také označení **PIM-SSM**. Pro nasazení je dále nutné zajistit fungování s přístupovým mechanismem IGMPv3, který podporuje filtrování zdrojů. Vzhledem k tomu, že implementace SSM závisí výhradně na budování nejkratších distribučních stromů *Shortest Path Trees* (SPT) a naopak eliminuje sdílené *Rendezvous Point Trees* (RPT), příjemce musí nejdříve získat specifickou adresu zdroje tak, aby mohl splnit notaci (S,G) nejkratšího distribučního stromu. Tuto získává většinou externě, např. z webové stránky nebo emailu. Příjemce tedy sám od sebe informuje svůj *first-hop router*, který okamžitě provádí *RPF check* a buduje tak nejkratší distribuční strom směrem k multicastovému zdroji.

SSM využívá vlastní multicastový IP rozsah 232.0.0.0/8, kde je zcela zakázáno vytváření sdílených distribučních stromů. SSM tak může fungovat i souběžně s klasickým ASM.

1.3.4 DVMRP, MOSPF a BGMP

Jedná se o multicastové nadstavby původně unicastových směrovacích protokolů. Vývoj většiny z nich ustal ještě před rokem 2000 [11], před rozvojem protokolu PIM. Nejúspěšnějším ze zmíněné trojice se stal BGMP, jehož specifikace byla upravena naposledy v roce 2004. Důvodem bylo podstatně lepší škálování při využití BGMP pro směrování multicastů než u jeho tehdejších konkurentů. Stejně jako BGP, BGMP využívá TCP jako svůj transportní protokol, eliminuje tak potřebu implementace fragmentování zprávy, přeposílání a potvrzování, jelikož tyto jsou již zahrnuty v TCP. Nová spojení jsou navazována na port 264, který je odlišný od původního BGP, které operuje na portu 179. Provoz obou protokolů je tak jasně odlišitelný. *KeepAlive* zprávy jsou zasílány periodicky stejně jako v případě BGP, notifikační zprávy jsou zasílány okamžitě po výskytu události.

Základní myšlenkou BGMP bylo směrování nikoliv mezi samotnými routery, ale mezi tzv. multicastovými doménami – každá popsána většinou jedním autonomním systémem - které mohly být uvnitř dále směrovány libovolnými multicastovými směrovacími protokoly. BGMP tedy pouze poskytuje další vrstvu pro vzájemné propojení těchto domén a výběr nejlepší cesty mezi nimi.

V praxi ale dochází k takovému propojení více nezávislých multicastových domén pouze výjimečně, a proto jsou BGMP i další výše zmíněné multicastové směrovací protokoly v dnešním internetu využívány pouze zřídka, a to pouze z historických důvodů.

1.4 Multicast na přístupových zařízeních

Zatímco mezi routery je multicastový provoz směrován multicastovými směrovacími protokoly, pro distribuci a správu vícesměrového vysílání na lokální síti je využít *Internet Group Management Protocol*.

1.4.1 Internet Group Management Protocol (IGMP)

IGMP je jednoduchý síťový protokol, který zajišťuje provoz vícesměrového vysílání na jedné síťové (broadcastové) doméně.

Stanice jej využívají k přihlášení k odběru multicastových paketů z dané multicastové skupiny a jejich následný příjem, odhlášení z dané multicastové skupiny, apod. Router zase k udržování stavu posluchačů jednotlivých multicastových skupin na svých síťových rozhraních.

Existují celkem tři verze protokolu IGMP:

- IGMPv1 (RFC 1112)
- IGMPv2 (RFC 2236)
- IGMPv3 (RFC 3376, RFC 4604)

IGMPv2 účastníkovi dává možnost aktivně opustit skupinu (*IGMP Leave*). Ukončení členství ve skupině je totiž v případě IGMPv1 realizováno jako vypršení časového limitu členství (standardně tři minuty) poté, kdy účastník přestane reagovat na *IGMP Query* lokálního routeru.

IGMPv3 [12] pak přidává možnost filtrování zdrojů, tedy možnost přihlásit se k odběru pouze ke specifickým zdrojovým adresám, jak to vyžaduje *Source-Specific Multicast*.

Přestože verze 3 úspěšně prošla standardizací a je zpětně kompatibilní s předchozími verzemi, její použití v praxi stále není obvyklé.

Cisco proprietární variantou IGMP je pak *Cisco Group Management Protocol (CGMP)*.

Typy zpráv

IGMP Report (*Host Membership Report*) je zpráva, kterou odesílá stanice v případě zájmu připojit se k dalším odběratelům v příslušné multicastové skupině. Cílovou IP adresou je přitom adresa multicastové skupiny, zdrojová IP je unicastová adresa daného účastníka.

IGMP report je také odesílán stanicí jako reakce na IGMP Query, kterou se router dotazuje svých stanic, zda trvá jejich zájem odebírat data pro příslušnou multicastovou skupinu.

V závislosti na implementaci účastník odesílá i duplicitní IGMP Report zpožděný v čase. Snaží se zvýšit pravděpodobnost doručení principiálně nespolehlivého IP datagramu.

Zprávou **IGMP Query** (*Host Membership Query*) směrovač žádá periodicky (standardně 60 sekund) účastníky na síti, aby oznámily aktuální požadavky pro přiřazení k multicastovým skupinám. Vzhledem k tomu, že zpráva odchází na multicastovou adresu 224.0.0.1, obdrží ji všichni účastníci na síti.

Odpověď od účastníků na síti nepřichází okamžitě, ba dokonce nemusí ani dorazit od všech příjemců. Aby nedocházelo k periodickému zahlcování sítě, účastníci na *IGMP Query* odpovídají vždy s náhodným zpožděním. Pokud účastník obdrží *IGMP Report* od jiného účastníka na síti pro multicastovou skupinu, kterou také odebírá, nevyšle svůj zpožděný *IGMP Report* vůbec – předpokládá totiž, že stejnou informaci už získal i router, proto není nutné jej dále informovat o tom, že na síti existují příjemci dané multicastové skupiny.

```

root@sonda2:~# tcpdump -i eth0.1559 -n ip proto 2
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0.1559, link-type EN10MB (Ethernet), capture size 96 bytes
19:02:21.893318 IP 172.17.240.1 > 224.0.0.1: igmp query v2
19:02:31.222062 IP 172.17.240.9 > 239.250.1.1: igmp v2 report 239.250.1.1
19:03:21.883376 IP 172.17.240.1 > 224.0.0.1: igmp query v2
19:03:28.497873 IP 172.17.240.9 > 239.250.1.1: igmp v2 report 239.250.1.1
^C
4 packets captured
4 packets received by filter
0 packets dropped by kernel

```

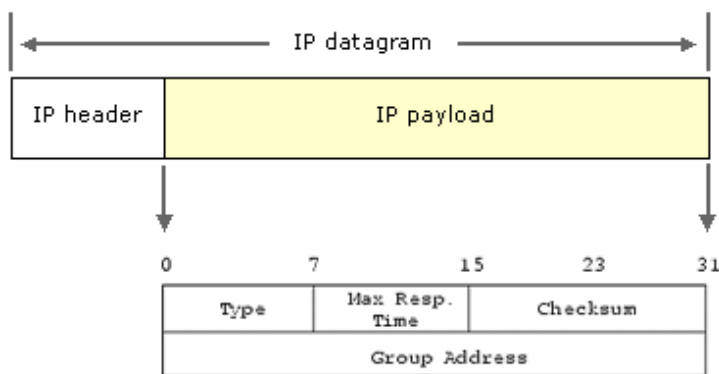
obr. 12 - zachycená zpráva IP *IGMP Query* a *IGMP Report*

Obecný dotaz *IGMP Query* je tzv. *General Query* – pole *Group Address* je nastaveno na nulu a odpoví na něj tedy nutně všichni účastníci na síti. Další variantou je tzv. *Group-Specific Query* - pole *Group Address* je nastaveno na žádanou adresu multicastové skupiny a odpovídají na ni tedy pouze účastníci z této konkrétní skupiny.

Zpráva **IGMP Leave** (*Leave group*), kterou přinesla druhá verze protokolu IGMP, umožňuje účastníkovi aktivně zrušit členství v dané multicastové skupině. Zpráva je adresovaná všem směrovačům na místní síti, tedy multicastové skupině 224.0.0.2. Argumentem je adresa skupiny, kterou chce účastník opustit.

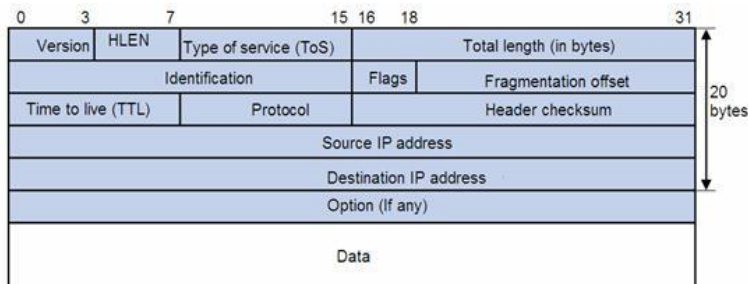
Formát

Typově jsou IGMP zprávy, podobně jako ICMP zprávy, obyčejné IP pakety bez transportní vrstvy. Níže je uvedena struktura zprávy IGMPv2, jelikož bude využívána v praktické části.



obr. 13 - IP datagram s IGMP zprávou [13]

- Type
 - Membership Query
 - Membership Report v1
 - Membership Report v2
 - Leave Group
- Max Resp Time
 - doba odezvy na Query
- Group Address
 - např. 239.250.1.1

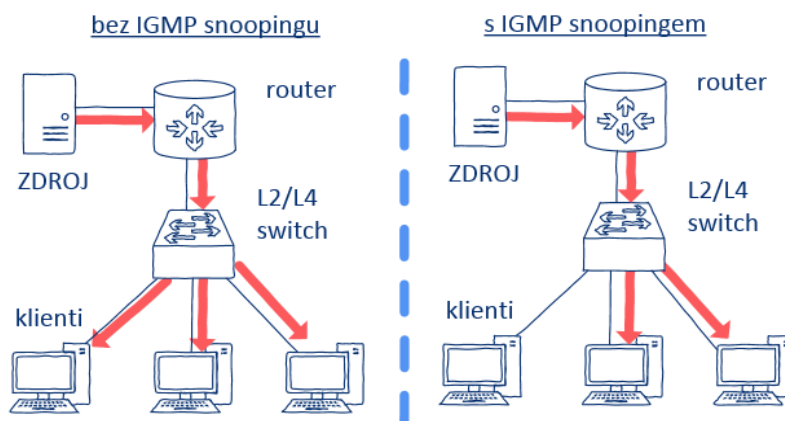


obr. 14 - obecná hlavička IP datagramu [14]

Délka IGMP zprávy je tedy pouze 8 B. Spolu s IP hlavičkou je to pak celkem 28 B. Time to live (TTL) hodnota je přitom nastavena vždy na 1 – IGMP zpráva tedy nikdy neopustí lokální síť, na které byla vytvořena.

IGMP Snooping

IGMP snooping, definovaný v RFC 4903 a 4541, je mechanismus využívaný především síťovými zařízeními na druhé vrstvě ISO/OSI modelu, tedy přepínači. Mohou díky němu úspěšně čelit nekontrolovanému šíření multicastu do všech portů přepínače, podobně jako linkový broadcast.



obr. 15 - IGMP snooping

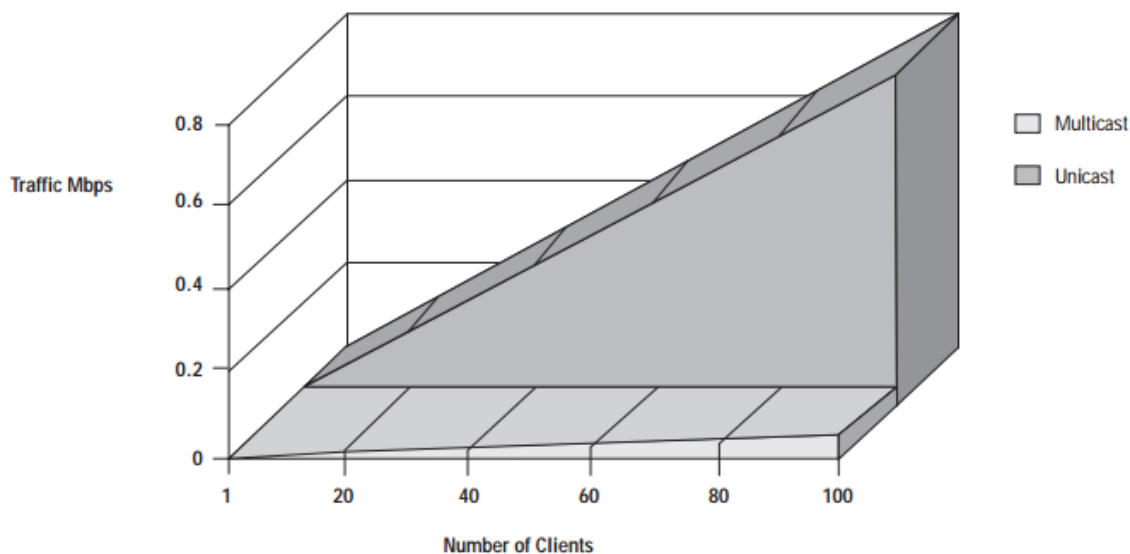
Problém tkví v přímém mapování MAC adres na multicastové IP adresy. Zatímco v případě unicastu je MAC adresa příjemce rámce získána dynamicky přes ARP protokol, v případě multicastu rámec míří na staticky přidělenou MAC adresu, kterou přepínač ve své tabulce nemůže vidět. Proto se k příchozímu rámci staví jako tzv. *unknown-unicastu*. Ten se standardně odbaví prvotním přeposláním na všechny porty, ovšem ani poté se přepínač kýženou MAC adresu na žádném ze svých portů nenaučí. Každý další příchozí multicastový datagram je i potom odbaven jako *unknown-unicastu*, tedy je prakticky replikován na všechny aktivní porty přepínače jako broadcast.

IGMP snooping tento problém řeší odchyťáváním a analýzou průchozích *IGMP Join* a *IGMP Leave* paketů. Vzhledem k tomu, že se skutečně jedná o pakety, přepínače podporující IGMP snooping už musí být schopny pracovat i s třetí vrstvou ISO/OSI modelu, byť pouze okrajově. Z informací obsažených v těchto zprávách mohou přepínače budovat seznam rozhraní a multicastových skupin jim náležících.

V RFC jsou definovány i další implementace IGMP snoopingu jako *Proxy reporting* nebo *IGMP Querier*. První agreguje odchozí IGMP Report zprávy a snižuje tak zátěž sítě odstraněním duplicitních zpráv zasílaných routeru. Druhá umožňuje suplovat roli multicastového Querier směrovače a rozesílat *IGMP Queries* na VLAN bez síťového rozhraní daného směrovače.

1.5 Aplikace a rozšíření

Vícesměrové vysílání najde uplatnění všude tam, kde existuje vysoký počet příjemců, kteří odebírají velké množství dat podobného charakteru. Vzhledem k podstatně lepšímu škálování při přenosu velkých datových objemů by měl být multicast upřednostňován před unicastem při realizaci IPTV, hromadné distribuci multimediálního obsahu nebo realizaci CDN (*Content Delivery Network*) sítí.



obr. 16 - grafické znázornění škálování multicastu ve srovnání s unicastem [15]

Využití nespolehlivých UDP datagramů přitom pro multimediální aplikace přenášející audio a video obsah nepředstavuje žádnou zásadní překážku. Kvalita telekomunikačních sítí navíc již dnes umožňuje přenášet velké objemy multimediálních dat, aniž by bylo třeba se spoléhat na opravné mechanismy, které unicast ze své podstaty nabízí.

I přesto všechno ale rozšíření vícesměrového vysílání na internetu dalece zaostává za nadějami, které do něj odborná veřejnost vkládala před již více než dvaceti let, kdy byl multicast poprvé oficiálně definován. Vinu lze hledat v samotné decentralizaci internetu a absenci jednotné správy, která by se o rozšíření multicastu zasadila a koordinovala jej. Aplikace se tedy většinou omezují na geograficky izolovaná území pod jednotnou správou, kde multicast účinně pomáhá snižovat nároky na přenosovou kapacitu. Na místní síti je i pro běžný unicastový provoz významný multicastový IP rozsah 224.0.0.0/24, tedy tzv. *Local Network Control Block*. Na těchto adresách běžně komunikují i unicastové směrovací protokoly a služby jako RIPv2 (224.0.0.9), OSPF (224.0.0.5) nebo služba VRRP (Virtual Router Redundancy Protocol) (224.0.0.18).

2 OSTATNÍ ZPŮSOBY VYSÍLÁNÍ

V prostředí internetových sítí se ale využívají i další techniky k doručení informace mezi jednotlivými prvky sítě. Pokud bychom na probíhající síťový provoz (*obr. 1*) nahlíželi skrze síťovou vrstvu, pravděpodobně nejčastěji by v něm byly zastoupeny datagramy s jediným cílem, tedy datagramy zaslané jako **unicast**. Další v pořadí by byly nejspíše datagramy zaslané všem hostům na lokální síti, tedy **broadcasty**. Tyto jsou přítomny na jakékoliv síti, protože jejich základní funkce je na unicastovém a broadcastovém zasílání datagramů závisí. V závislosti na dalších aplikacích se setkáme i s **multicastem**, tedy datagramy určenými dané skupině příjemců. Základní trojici doplňuje **anycast**, který z definované skupiny příjemců vybírá nejbližšího člena a kombinuje tak unicastová zasílání s multicastovým. Dále také některé zdroje uvádí tzv. **geocast**, využívaný v některých mobilních aplikacích.

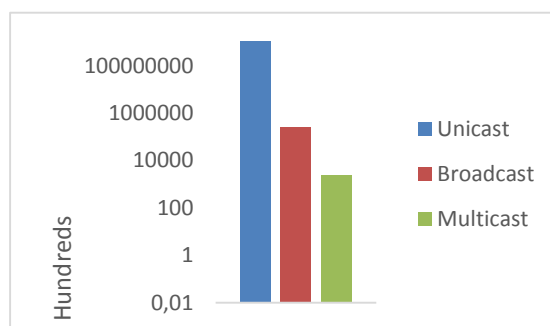
```
Peak value of input: 416937269 bytes/sec, at 2013-01-09 23:49:15
Peak value of output: 151423460 bytes/sec, at 2013-01-09 21:10:59
Last 300 seconds input: 117236 packets/sec 116642695 bytes/sec 10%
Last 300 seconds output: 99746 packets/sec 78847270 bytes/sec 6%
Input (total): 1245741116659 packets, 1363477051195361 bytes
968860868009 unicasts, 57054139 broadcasts, 276823194408 multicasts, 0 pauses
Input (normal): 1245741116556 packets, - bytes

968860868009 unicasts, 57054139 broadcasts, 276823194408 multicasts, 0 pauses

Input: 79 input errors, 0 runts, 0 giants, 0 throttles
       79 CRC, 0 frame, - overruns, 0 aborts
```

obr. 17 – statistika přijatých paketů na portu switchu

Poslední dva uvedené způsoby zasílání datagramů ale nemají přidělen vlastní adresní prostor, proto je nelze snadno identifikovat. Obrázek výše je výpis statistiky odeslaných a přijatých paketů na portu L3 switchu, kterým prochází agregovaný provoz několika stovek běžných uživatelů internetu. Simuluje tedy průměrný internetový provoz.

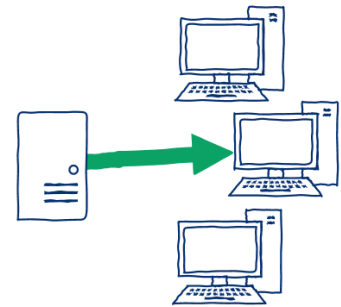


obr. 18 – poměr paketů na internetu

Celkový počet odeslaných paketů je v grafu (*obr. č. 18*) možné porovnat jen díky logaritmické ose. Počet unicastových paketů totiž mnohonásobně převyšuje oba ostatní způsoby adresace. Unicast je v dnešní době využíván prakticky v každé síťové aplikaci.

2.1 Unicast

Jak už bylo uvedeno výše, unicast v soudobých sítích představuje majoritní část probíhajícího síťového provozu. Prakticky všechny z internetových protokolů aplikační vrstvy (FTP, HTTP, POP3, Telnet,...) navazují spojení s konkrétním hostem, a tedy posílají data jako unicast.



obr. 19 – unicast

2.1.1 Adresování

Pro adresování se v praxi využívá pouze část z celkového 32bitového IPv4 rozsahu. Část je rezervována pro privátní použití, další rozsahy jsou historicky vázány na konkrétní společnosti, popř. experimentální využití a jiné typy adresování.

2.1.2 Směrování

Mezi unicastové směrovací protokoly patří RIP, OSPF, BGP nebo Cisco proprietární EIGRP, i když některé z nich, např. OSPF, šíří informace o změnách topologie sítě pakety adresovanými jako *multicast*.

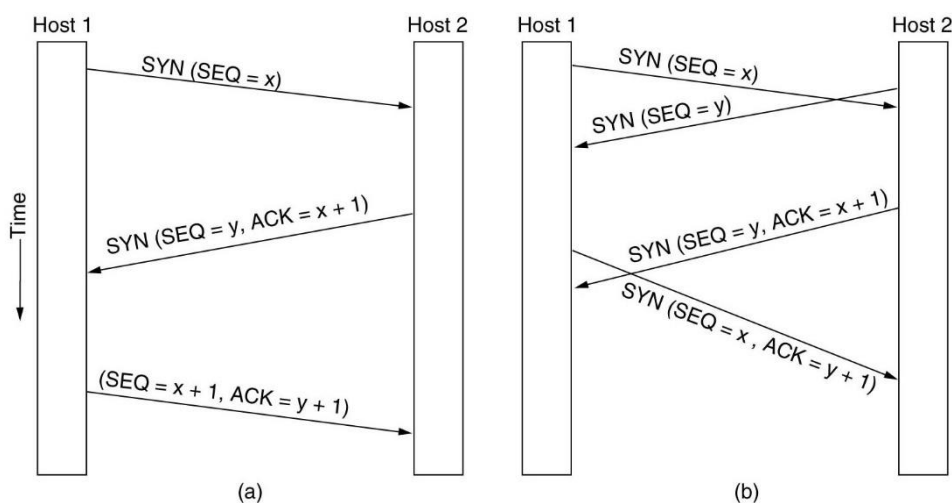
2.1.3 TCP/UDP

Z přenosových protokolů transportní vrstvy využívají výhod unicastového adresování jak TCP, tak i UDP, avšak pouze TCP exklusivně – udržet spojení je možné pouze s jediným příjemcem zprávy. Zaslání datagramů UDP pouze mírně rozšiřuje schopnosti vrstvy IP o možnost komunikace na více portech jedné IP adresy. Bezstavovost a úsporná hlavička přináší výrazné snížení režie i latence, avšak zavádí do přenosu asymetrii, nemožnost regulovat datový tok nebo zaslat opětovně ztracený paket.

2.1.4 TCP Three-Way Handshake

Už ze své podstaty je tedy unicastový přenos paketů pevně spojen s TCP, tedy nositelem funkce čtvrté vrstvy ISO/OSI modelu. Jakýkoliv jiný způsob vysílání, tedy skupině hostů (v případě multicastu) nebo všem hostům (v případě broadcastu), nedokáže transportní vrstvě poskytnout relevantní zpětnou informaci o doručení paketu, kterou transportní vrstva vyžaduje k udržení spojení.

Při jeho navazování probíhá tzv. “TCP Three-Way Handshake”, tedy vzájemná komunikace obou stran přenosu.



obr. 20 – unicastový TCP Three-Way Handshake

V prvním kroku dochází k otevření socketu a odeslání inicializačního SYN (synchronized) paketu, nastavení časovače pro opětovné poslání (retransmission timer) a čekání na návrat odpovědi od vzdáleného socketu. SYN paket obsahuje prázdný TCP segment, nastavený příznak SYN v TCP hlavičce, dále informuje o velikosti okna/bufferu (Win) a maximální velikosti segmentu (MSS) k předcházení fragmentování.

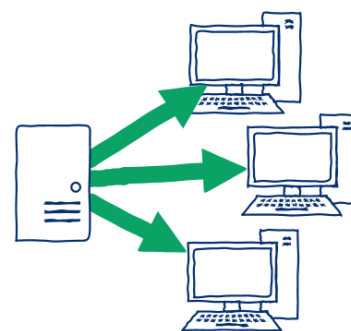
V druhém kroku dochází k vytvoření procesu a extrakci obsahu samotného paketu na síťové vrstvě, tedy navázání spojení. Nově vytvořený proces odesílá SYN-ACK paket. Při tom dochází k opětovnému nastavení časovače (retransmission timer) stejně jako v předchozím kroku. Odesílatel generuje vlastní sekvenční číslo (SEQ), dále zvyšuje původní příchozí sekvenční číslo o jedna a posílá protistraně jako Acknowledgment number (ACK) zpět.

Po přijetí paketu synchronizace (SYN-ACK) posílá příjemce zpět paket potvrzující přijetí potvrzení. Paket má nastaven pouze příznak ACK, nikoliv SYN. V případě přijetí protistranou začíná komunikace.

2.2 Broadcast

Všesměrové vysílání hraje v prostředí počítačových sítí poměrně kontroverzní roli. Na jedné straně je zapotřebí k zajištění samotného fungování sítě, na straně druhé může při nevhodném využití její fungování výrazně omezit až zastavit úplně.

Broadcastový datagram samozřejmě bez výjimky využívá UDP socket transportní vrstvy ISO/OSI modelu, jde tedy o bezstavové vysílání.



obr. 21 – broadcast

Pro adresování na třetí vrstvě je využita tzv. IP broadcastová adresa, která má dvě podoby:

- **omezený broadcast** (*limited broadcast*)
 paket s cílovou adresou 255.255.255.255, který obdrží všichni ostatní účastníci na síti odesílatele. Tento druh broadcastu není směrován routery, proto obsáhne pouze místní síť.
- **řízený broadcast** (*directed broadcast*)
 paket s cílovou adresou závislou na cílové podsíti, např. 192.168.255.255 pro podsít' 192.168.0.0/16. Routery na síti mohou být nastaveny k přeposílání řízeného broadcastu.

Popis výpočtu adresy řízeného broadcastu pro 10.64.172.131/26:

Krok č.1 - binární invertování masky podsítě

maska podsítě:	11111111 11111111 11111111 11000000
invertovaná maska podsítě:	00000000 00000000 00000000 00111111

Krok č.2 - logický součet IP adresy a invertované masky podsítě

IP adresa binárně:	00001010 01000000 10101100 10000011
invertovaná maska podsítě:	00000000 00000000 00000000 00111111
broadcastová IP adresa:	00001010 01000000 10101100 11111111

(tedy dekadicky 10.64.172.191)

Při průchodu zprávy referenčním ISO/OSI modelem je každý nově vzniklý rámec adresován na tzv. broadcastovou adresu linkové vrstvy (*Data link layer broadcast address*), konkrétně FF:FF:FF:FF:FF:FF. Typických aplikací broadcastu na síti je hned několik, nejčastěji je to realizace protokolů aplikační vrstvy jako ARP/RARP, DHCP nebo jako komunikační médium starších směrovacích protokolů odvozených od protokolu RIP.

2.2.1 ARP - Address Resolution Protocol

V případě komunikace mezi účastníky na stejné podsíti je ARP protokol využíván k namapování IPv4 adresy třetí vrstvy na linkovou adresu rozhraní MAC druhé vrstvy. Mechanismus přitom využívá broadcastového dotazu, tzv. *ARP Request*, k získání MAC adresy fyzického rozhraní příjemce. Dotaz tedy obdrží všichni účastníci na síti, odpověď od účastníka s hledanou adresou (tzv. *ARP Reply*) ale už přichází tazateli nazpět jako unicast, protože jeho adresa je uvedena v prvotním dotazu.

```
router:~# tcpdump -i eth1.1124 -nv arp
...
18:05:29.602012 ARP, Ethernet (len 6),
IPv4 (len 4), Request who-has 10.24.0.254 tell 10.24.3.109, length 46
18:05:29.602028 ARP, Ethernet (len 6),
IPv4 (len 4), Reply 10.24.0.254 is-at 00:25:90:68:83:b6, length 28
...
18:05:30.432641 ARP, Ethernet (len 6),
IPv4 (len 4), Request who-has 10.24.7.10 tell 10.24.0.1, length 28
18:05:30.433081 ARP, Ethernet (len 6),
IPv4 (len 4), Reply 10.24.7.10 is-at 00:02:72:6a:9d:67, length 46
```

obr. 22 – zachycení *ARP Request* a *ARP Reply*

2.2.2 DHCP - Dynamic Host Configuration Protocol

Dynamické přidělování IP adres a dalších údajů nutných pro komunikaci účastníka na síti (výchozí brána, DNS servery, maska podsítě, ...) opět z velké části závisí na prvotním broadcastovém požadavku (*DHCP discover message*), kterým se nově připojený účastník táže všech ostatních hostů na síti, zdali mu nedokáží poskytnout potřebné informace.

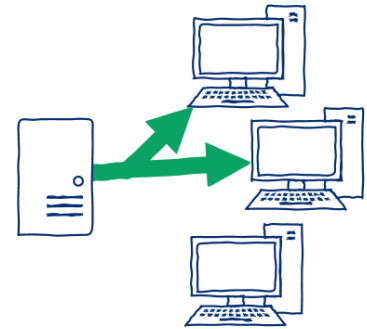
2.2.3 Broadcast storm

Stinnou stránkou broadcastového vysílání je ale tzv. broadcastová bouře, tedy nekontrolovaný nárůst broadcastů na síti, který vzniká nejčastěji fyzickým, ale i logickým zakruhováním sítě. V té chvíli jednak selhávají všechny základní přepínací mechanismy síťových prvků, jednak se podobné závady jen těžko analyzují a opravují.

2.3 Anycast

Podobně jako v případě všesměrového či vícesměrového vysílání anycast adresuje datagram více příjemcům. Dorazí však pouze jednomu z nich, a to zpravidla nejbližšímu účastníkovi [16].

Na rozdíl od unicastu nebo multicastu nemá anycast žádný přidělený IP rozsah – všichni příjemci datagramu mají stejnou IP adresu z unicastového IP rozsahu. Veškerá logika se odehrává na BGP routerech.



obr. 23 – anycast

Ve většině případů je anycastové adresování realizováno n-násobnou dynamickou propagací stejného IP rozsahu směrovacím protokolem BGP. V tomto případě router přijaté anycastové UDP datagramy směřuje jednoduše na rozhraní s nejnižší metrikou k požadovanému cíli.

Anycast je běžně využíván při realizaci DNS, kde je klíčová co nejrychlejší odezva serveru (např. známá adresa Google DNS 8.8.8.8, viz. příloha). Další výhodou je vysoká míra redundance – na jedné IP adrese odpovídá i několik serverů umístěných v různých geograficky oddělených oblastech.

V poslední době se ale začínají objevovat aplikace využívající směrování TCP paketů i jako Anycast, konkrétně moderní CDN (Content Delivery Network) sítě, umožňující udržení navázaného TCP spojení i v případě selhání preferovaného zdroje v průběhu přenosu. Toto ale nelze provést beze změny jak na straně klienta (ošetření výjimek po přijetí chybného TCP soketu na aplikační vrstvě), tak i na straně serverů (pouze výkonnostní optimalizace).

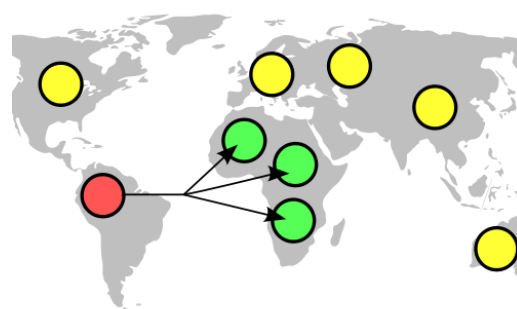
V neposlední řadě anycastové směrování účinně potlačuje možnost hromadného útoku na službu – zátěž je totiž rovnoměrně rozprostřena na různé fyzické stroje podle geografické polohy každého z účastníků.

2.4 Geocast

Hlavní myšlenka Geocastu spočívá v uvedení GPS souřadnic do směrovacího schématu a jejich funkčního propojení s IP adresou [17].

Původně uvažovanou možností realizace byla pouhá nadstavba v aplikační vrstvě, tedy v podstatě tunelování standardního IP provozu a následné broadcastování zpráv příjemcům v geografické oblasti označené polygonem s vrcholy s danými GPS souřadnicemi. Později se ale Geocast vyvinul ve specifickou podmnožinu vícesměrového vysílání, kde je skupina příjemců identifikována jejich společnou geografickou polohou.

Původně byl Geocast designován jako pomůcka k adresování geograficky oddělené skupině příjemců, např. při nastalé přírodní katastrofě nebo jiné mimořádné události. Koncept ale nepřekročil hranici návrhu a v praxi se nikdy neprosadil.



obr. 24 – geocast [18]

II. PRAKTICKÁ ČÁST

3 VYUŽITÍ MULTICASTU PRO PŘENOSY IPTV

Přestože je vícesměrové vysílání velmi efektivním systémem síťové komunikace a jeho základy byly položeny už v roce 1986, na reálné využití čekal velmi dlouho. To přichází zdá se až nyní s rozvojem IPTV a Video on Demand (VoD) služeb, které přesně zapadají do *one-to-many* konceptu vícesměrové komunikace. Kvalitativní nároky na služby podobného charakteru jsou ale řádově vyšší než na běžné IP služby. Proto byl vyvinut i objektivní systém na jejich měření.

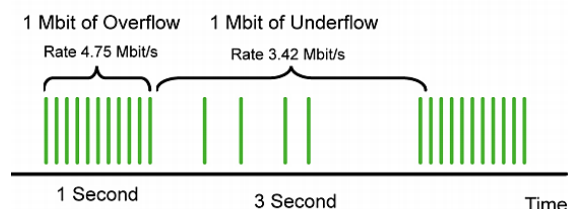
3.1 Media Delivery Index (MDI)

MDI představuje řadu nástrojů, které mohou být použity k objektivnímu vyhodnocení kvality přenášovaných video streamů z pohledu přenosových chyb, které mohou na síti v době přenosu nastat. Media Delivery Index bývá zobrazován jako dvě čísla oddělená dvojtečkou:

Delay Factor (DF): Media Loss rate (MLR

3.1.1 Delay Factor (DF)

V případě, že aktuální parametry sítě nedovolují příjemci získávat souvislý datový tok stejnou rychlostí jakou je šířen zdrojem do sítě, příchozí pakety musí být pro plynulé přehrávání ukládány do vyrovnávací paměti. Čím větší je rozptyl zpoždění paketu (jitter), tím větší vyrovnávací paměť (buffer) je potřeba. Se zvyšujícím se bufferem ale také může růst zpoždění vysílání. Nastat ale může i opačná situace, kdy jsou data příjemci dodávána rychleji, než je dokáže zpracovat.



obr. 25 - příjem zpožděných paketů [19]

DF je definován jako čas (v ms), který je zapotřebí k eliminaci tohoto zpoždění. Akceptovatelné rozmezí leží mezi 9-50 ms [20].

$$X = |\text{Bytes_received} - \text{Bytes_drained}| \qquad DF = \frac{[\max(X) - \min(X)]}{\text{přenosová rychlost média}}$$

- *Bytes_received*
příchozí data do vyrovnávací paměti
 - *Bytes_drained*
data odčerpaná z vyrovnávací paměti
- *min(x), max(x)*
hraniční hodnoty měřeného intervalu
 - *přenosová rychlost média*
udaná v B/s

3.1.2 Media Loss Rate (MLR)

Hodnota MLR je definována jako počet ztracených paketů, nebo počet paketů přijatých mimo pořadí. Hodnota MLR je vypočítána jako rozdíl počtu očekávaných a skutečně přijatých paketů v měřeném intervalu.

$$MLR = \frac{(\text{packet_expected} - \text{packets_received})}{\text{interval}}$$

služba	max. MLR
SDTV	0,004
VoD	0,004
HDTV	0,0005

tab. 1 - max. doporučené hodnoty MLR

Tabulka ukazuje přehled maximálních doporučených hodnot MLR pro zachování přijatelné kvality přenosu různých IPTV služeb. Pro HDTV přenosy to v absolutním vyčíslení znamená pět po sobě jdoucích ztracených paketů za čtyři hodiny vysílání. Hodnota DF:MLR tak představuje objektivní a porovnatelné měřítko k posuzování kvality přenosu služeb spojených s IPTV.

3.2 Kvalitativní požadavky na přenos IPTV

Mezi základní kvalitativní požadavky přenosu tedy patří propustnost, dále také latence, ztrátovost a změna pořadí paketů.

3.2.1 Propustnost

Obrazová data jsou z historických důvodů přenášena ve dvou rozlišeních, a to 576 řádků (PAL) a 1 024 řádků (HD). Zatímco video s kvalitou PAL je standardně přenášeno jako 50i, tedy 50 půlsnímků, HD video je přenášeno s plnými 50ti snímky. K obrazovým datům je dále nutné připočítat i zvukovou stopu. V případě PAL signálu jsou to zpravidla dva kanály v režimu stereo, v případě HD signálu pak jako vícekanálové Dolby Digital (AC3). Výpočet maximální potřebné propustnosti linky by tedy byl následující.

Pro PAL

$$| \quad 720 \text{ px} \times 576 \text{ řádků} \times 25 \text{ snímků} = 10\,368\,000 \text{ bit} / 1\,024^2 = \sim 10 \text{ Mbps}$$

Pro HD

$$| \quad 1920 \text{ px} \times 1024 \text{ řádků} \times 50 \text{ snímků} = 98\,304\,000 \text{ bit} / 1\,024^2 = \sim 94 \text{ Mbps}$$

Dále je také nutno přičíst kapacitní nároky zvukové stopy.

Ke snížení datové náročnosti jsou jak obrazová, tak i zvuková data dále komprimována ztrátovými kompresními algoritmy. V případě HD signálu standardem MPEG-4, v případě PAL signálu nejčastěji standardem MPEG-2.

Výsledný datový tok závisí na mnoha proměnných - individuální nastavení poskytovatele streamu, způsob šíření, charakter přenášeného obrazu, atd. Pro signál v rozlišení PAL se ale nejčastěji uvádí průměr **~4 Mbps** na kanál, v případě HD rozlišení pak **~10 Mbps**.

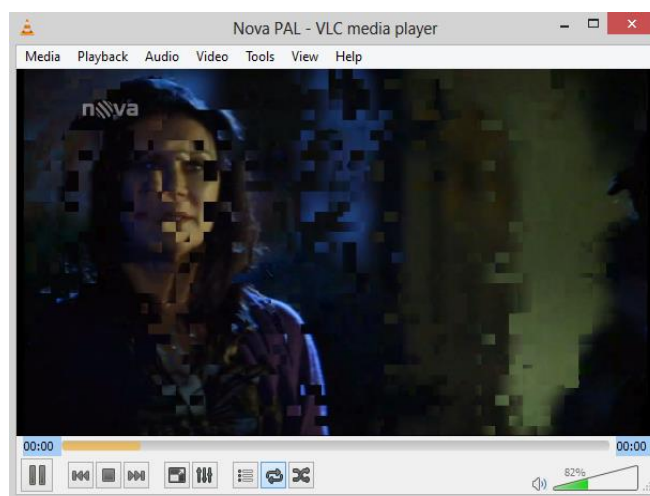
Na televizním trhu přitom operují stovky takovýchto stanic. V popisovaném případě představuje souhrnný datový tok všech televizních stanic asi **700Mbps**, tedy datový provoz, který by bez vícesměrového vysílání nikdy nebylo možno odbavit.

3.2.2 Latence

Přítomnost vyšší odezvy má vliv také na rychlost přepínání kanálů. V prostředí vícesměrového vysílání bývá toto zpoždění nazýváno také jako *igmp latency*, tedy čas potřebný k přihlášení a odběru dat nové multicastové skupiny. Nejde přitom pouze o skutečnou rychlost doručení požadavku na změnu kanálu ke streamovacímu serveru a zpět. Roli hraje také charakter kódovaného video streamu, rozložení a výskyt klíčových snímků apod.

3.2.3 Ztrátovost

Už z principu své funkce, tedy zasílání UDP datagramů, je jasné, že multicastové vysílání bude na ztrátu paketů mnohem více citlivé než vysílání unicastové. Zatímco v případě unicastového TCP spojení je každý přenesený byte označen sekvenčním číslem, jehož příjem je druhou stranou aktivně potvrzován ACK zprávami, UDP datagramy tuto vlastnost nemají. Takové chování sice podstatně snižuje režii přenosu (UDP hlavička 8B vs. minimální TCP hlavička 20B + další režie potřebná pro navázání a udržení spojení), na druhou stranu znemožňuje využití technik TCP spolehlivého doručování, především mechanismů jako *Flow Control* a *Congestion Control*. Pokud je spolehlivé doručení vyžadováno, musí být implementováno na vyšší vrstvě ISO/OSI modelu (např. aplikační protokol pro přenos souborů TFTP).



obr. 26 - obrazové artefakty při ztrátě paketu

Absence vyrovnávací paměti jak na straně odesílatele, tak i příjemce, a tedy i zmíněných TCP mechanismů, které s touto pamětí pracují, výrazným způsobem zvyšuje nároky na kvalitu datové linky. Dekodér video streamů totiž nemá k dispozici data z vyrovnávací paměti a nemůže z nich tak složit výsledný obraz.

3.2.4 Změna pořadí paketů

Vlivem probíhající konvergence směrovacích protokolů na síti a replikaci UDP multicastových datagramů na síťových routerech může dojít i k tomu, že koncový uzel obdrží multicastový datagram buď duplikovaný nebo v nesprávném pořadí.

3.3 Quality of Service (QoS)

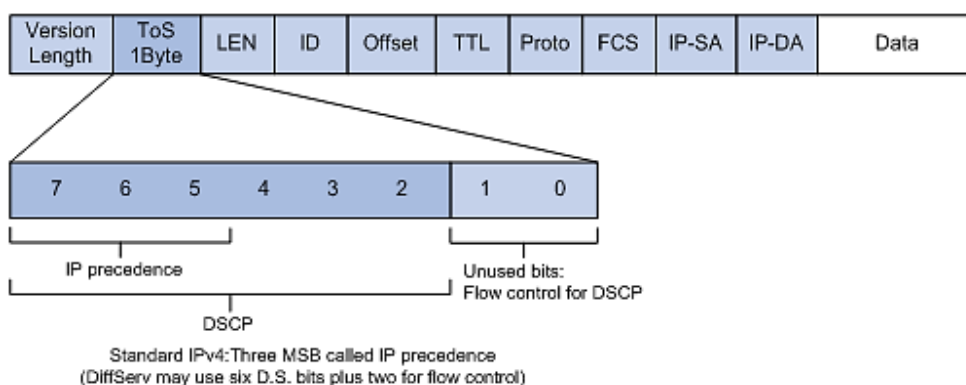
K potlačení všech výše zmíněných negativních vlivů se proto v IPTV sítích často využívá principu QoS, s jehož pomocí je možné prioritizovat kritický provoz (v tomto případě IPTV) a garantovat tak pro něj požadovanou šířku pásma na úkor služeb, u kterých není 100% dostupnost klíčová (typicky internetové protokoly jako http, smtp, apod.).

Prakticky celý internet byl původně postaven na tzv. principu *best effort policy*, který všem prvkům na síti, jejich uživatelům i procesům měří “stejným metrem”, a proto negarantuje žádnou *quality of service*.

V prostředí počítačových sítí je možné QoS aplikovat na síťové i linkové vrstvě ISO/OSI modelu. Na síťové vrstvě jako *DiffServ Code Points (DSCP)*, na linkové vrstvě pak technikou obecně známou jako *Class of Service (CoS)*. Obecně je přínos QoS nejmarkantnější především na místech, která představují úzká hrdla sítě, jako jsou např. bezdrátové spoje.

3.3.1 DiffServ Code Points (DSCP)

Differentiated services (nebo *DiffServ*, definováno v RFC 2474) představují v prostředí dnešního internetu nejčastější způsob implementace QoS. DSCP využívá 6 bitů v hlavičce IP paketu, původně označené jako ToS byte, dnes DS-field. Celkem tedy může nabývat 64 různých stavů.

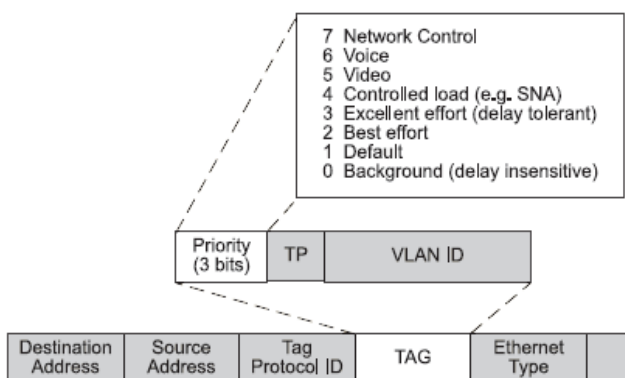


obr. 27 - DSCP bity v hlavičce IP paketu [21]

Vzhledem k tomu, že mapování různých druhů provozu do konkrétních hodnot DSCP je pouze doporučeno (RFC 4594), může se snadno vyskytnout problém různého mapování při přechodu paketu sítěmi pod rozdílnou technickou správou. V prostředí globálního internetu tedy původně navrženou funkci QoS nikdy nelze zaručit.

3.3.2 Class of Service (CoS)

Byť je *Class of Service* jedním z výsledků činnosti skupiny *IEEE P802.1p* a často se také jako 802.1p označuje, nejedná se o standard, spíše o techniku, která se samotná stala součástí standardu *IEEE 802.1q*. Jedná se o tři bity PCP (tzv. *Priority Code Point*), které jsou vloženy do 802.1q tagu rámců linkové vrstvy. Na druhé vrstvě ISO/OSI modelu lze tedy celkem rozlišit 2³ QoS stavů. Jejich rozdělení a definice je přitom pouze doporučeno. Administrátoři s nimi tedy mohou volně nakládat podle požadavků sítě. Vyšší číslo znamená vyšší prioritu provozu.



obr. 28 - 802.1q a CoS v ethernetovém rámci [22]

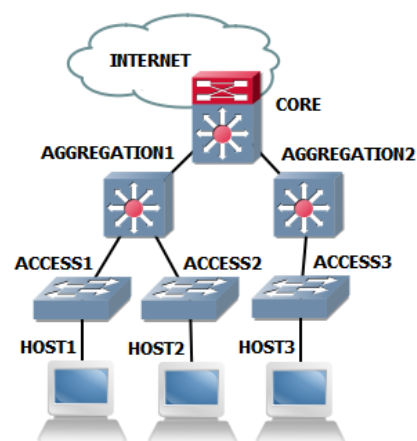
PCP	priorita	zkratka	typ přenosu
1	0	BK	pozadí
0	1	BE	best effort
2	2	EE	excellent effort
3	3	CA	kritické aplikace
4	4	VI	video, nižší priorita
5	5	VO	hlas, vyšší priorita
6	6	IC	internet control
7	7	NC	network control

tab. 2 - doporučení CoS hodnot

3.3.3 Mapování DSCP na CoS

Pokud je při reálném nasazení potřeba přenést nastavení QoS ze síťové (IP paketů) na linkovou (rámce) vrstvu ISO/OSI modelu pro prioritizaci provozu na síťových prepínačích, přichází na řadu mapování hodnot DSCP na CoS.

Vzhledem k tomu, že DSCP nabývá až 64 stavů, ale CoS pouze 8 stavů, nutně musí dojít k vytvoření skupin, které budou na linkové vrstvě zpracovávány se stejnou prioritou. Zdaleka ne všechny síťové prepínače ale umí pracovat s osmi prioritními frontami. Častější je podpora čtyř prioritních front, na fyzické vrstvě tedy dochází k dalšímu sjednocení priorit. V případě IPTV je tedy multicastový datagram na Core prvku podle příchozí VLAN "označován" příslušným DSCP už při vstupu do sítě poskytovatele. Switche na nižších vrstvách už tuto nastavenou hodnotu pouze akceptují a dále přeposílají.



obr. 29 - obecná topologie sítě

3.4 Platforma multicastové IPTV

3.4.1 Nangu.TV

Nangu.TV je český poskytovatel řešení IPTV a OTT (Over The Top), mnohými označovaný jako soudobý technologický standard. Jeho kompletní řešení pokrývá všechny aspekty potřebné k doručení multimediálního obsahu až ke klientským zařízením. Internetovému poskytovateli zprostředkovává propojení s poskytovateli video obsahu a poskytuje kompletní klientskou hw platformu.

3.4.2 Set top box

Motorola VIP1003 představuje koncové klientské zařízení, které se IGMP dotazy přihlašuje k odběru daných multicastových skupin. Každá ze skupin přitom představuje jeden televizní kanál. Set top box má díky upravenému firmwaru od společnosti Nangu.TV i další přidané funkce, včetně vzdálené správy a konfigurace (tzv. *provisioning*).



obr. 30 - Motorola VIP1003

3.4.3 Nelineární služby

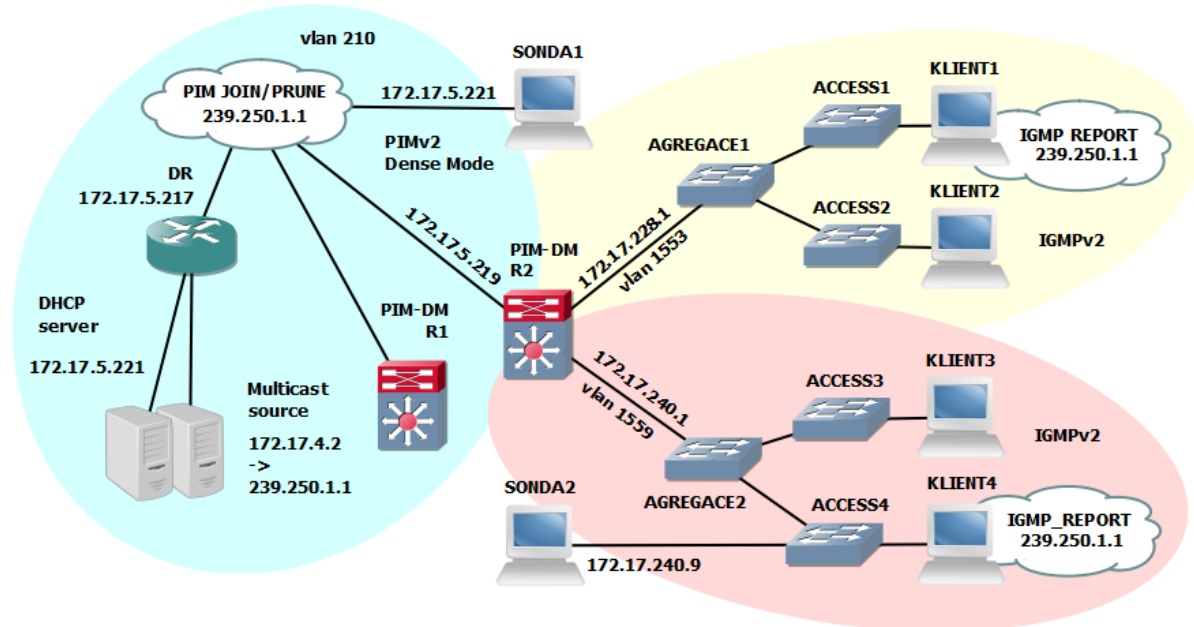
Jednou z přidaných funkcí jsou tzv. VoD (Video On Demand), tedy nelineární služby. Právě přítomnost obousměrné komunikace představuje nejvýraznější pokrok od minulých televizních platform založených ve své podstatě na broadcastování televizního signálu (DVB-x). Set top box např. umožňuje příjem v minulosti zaznamenaných programů v osobního archivu uživatele. Tato data jsou zaznamenána na diskových polích Nangu.TV a uživatelům jsou distribuována jako unicast.

3.4.4 Šifrování přenosu

Vzhledem k tomu, že IP multicast nezahrnuje podporu šifrování přenosu, podporuje hw platforma Nangu.TV i standardizovaná šifrování *Verimetric* a *SecureMedia*. Implementace je nutná ke splnění požadavků, které jsou na distributory vysílání kladeny vlastníky prémiového obsahu.

3.5 Síťová topologie

Topologie níže představuje velmi zjednodušený model síťové infrastruktury lokálního poskytovatele internetu, na které je služba multicastové IPTV provozována.



obr. 31 – zjednodušená topologie sítě

Celý model je rozdělen na tři oblasti.

- **modrá oblast** představuje PIM doménu, kde je multicastový provoz směřován mezi tzv. *PIM neighbors*, viz. dále.
- **žlutá oblast** představuje IGMP doménu, tedy přístupovou část, kde jsou multicastové datagramy přepínači s podporou *IGMP Snoopingu* efektivně doručovány jednotlivým posluchačům.
- **červená oblast** plní stejnou funkci jako žlutá oblast s tím rozdílem, že je oddělena jak geograficky, tak i odlišnou podsítí (VLAN 1553 vs. VLAN 1559).

Zastoupeny jsou nejen veškeré nutné části síťové infrastruktury, tedy vrstva *přístupová* (ACCESS), *agregační* (AGREGACE), ale i její klíčový prvek, tedy *hraniční router* celé sítě (CORE). Znázorněn je i *multicastový zdroj* dat se svým *Designated Routerem*, a v neposlední řadě i samotní *klienti* odebírající multicastové streamy.

Do topologie byly přidány i dvě *sondy* pro logování a analýzu průchozích dat.

3.5.1 PIM-DM doména

PIM doména je z větší části spravována dodavatelem multicastových streamů. Pod přímou správou poskytovatele internetu je pouze hraniční router jeho sítě (IP 172.17.5.219 na VLAN 210), který je zároveň sousedem (*PIM-neighbor*) Designated routeru multicastových zdrojů dodavatele (IP 172.17.5.217 na VLAN 210), viz. náčrt topologie sítě.

<pre><HP>display pim neighbor verbose Neighbor: 172.17.5.217 Interface: Vlan-interface210 Uptime: 4w:6d Expiry time: 00:01:35 DR Priority: 1 Generation ID: 0x7185D3B6 Holdtime: 105 s LAN delay: 500 ms Override interval: 2500 ms State refresh interval: 60 s Neighbor tracking: Disabled Bidirectional PIM: Disabled</pre>	<pre><HP>display pim neighbor verbose Neighbor: 172.17.5.218 Interface: Vlan-interface210 Uptime: 4w:6d Expiry time: 00:05:27 DR Priority: 1 Generation ID: 0x35F4E5EB Holdtime: 367 s LAN delay: 500 ms Override interval: 2500 ms Neighbor tracking: Disabled Bidirectional PIM: Disabled</pre>
---	---

obr. 322 - výpis PIM-neighbors na hraničním routeru

Druhý z PIM routerů na společné doméně (IP 172.17.5.218 na VLAN 210) je další z odběratelů multicastových streamů, který není pod správou lokálního poskytovatele internetu.

Oba tyto routery mají nastavenou stejnou hodnotu *DR Priority*. Standardně *Designated router* volí tzv. query router (*IGMP Querier*), který rozesílá *IGMP Queries* koncových uzlům. *PIM-DM* ve spojení s *IGMPv2/3* určuje *Query router* podle přidělené IP adresy, která je součástí *PIM Hello* paketů, které routery rozesílají v pravidelných intervalech (*PIM hello interval*, standardně 30 sekund) svým sousedům. Designated routerem je zvolen směrovač s nejvyšší IP adresou na daném segmentu sítě. Pokud nedojde k obnovení stavu (tzv. *PIM hello hold interval*, standardně 105 sekund), dojde ke zvolení nového DR.

V nastavení PIM routeru dalších odběratele chybí *State refresh interval*, proměnná určující, jak často dochází k obnově a prodloužení platnosti v minulosti odříznutých větví (*Prune*). Nemusí tak docházet k opakované záplavě PIM routerů na síti. Tato funkce je ale záměrně vypnutá na obou PIM routerech, k pravidelné záplavě tedy dochází (každých 180 sekund).

Oba PIM sousedi hraničního routeru dále nemají zapnutou podporu *PIM Source-specific* multicastu – PIM doména není nijak rozsáhlá, a tedy není zapotřebí dále zvyšovat bezpečnost provozu použitím PIM-SSM.

```

<HP>display pim interface Vlan-interface 210 verbose
Interface: Vlan-interface210, 172.17.5.219
PIM version: 2
PIM mode: Dense
PIM DR: 172.17.5.219 (local)
PIM DR Priority (configured): 2
PIM neighbor count: 2
PIM hello interval: 30 s
PIM LAN delay (negotiated): 500 ms
PIM LAN delay (configured): 500 ms
PIM override interval (negotiated): 2500 ms
PIM override interval (configured): 2500 ms
PIM neighbor tracking (negotiated): disabled
PIM neighbor tracking (configured): enabled
PIM generation ID: 0xD712D33A
PIM require generation ID: disabled
PIM hello hold interval: 105 s
PIM assert hold interval: 180 s
PIM triggered hello delay: 5 s
PIM J/P interval: 60 s
PIM J/P hold interval: 210 s
PIM state-refresh processing: disabled
PIM graft retry interval: 3 s
Number of routers on network not using DR priority: 0
Number of routers on network not using LAN delay: 0
Number of routers on network not using neighbor tracking: 2

```

obr. 33 - výpis PIM interface na vlan 210 hraničního routeru

```

<HP>display multicast rpf-info 172.17.4.2
RPF information about source 172.17.4.2:
  VPN instance: public net
  RPF interface: Vlan-interface210, RPF neighbor: 172.17.5.217
  Referenced route/mask: 172.17.4.0/24
  Referenced route type: unicast
  Route selection rule: preference-preferred
  Load splitting rule: disable

```

obr. 34 - zobrazení multicast reverse patch forwarding ke zdroji

Z výpisu je zřejmé, že multicastový zdroj dat je dostupný přes PIM souseda 172.17.5.217, tedy DR zdroje. Tento router je dostupný z hraničního routeru přes VLAN rozhraní 210. RPF-check byl proveden na základě informací z unicastové směrovací tabulky.

Pro spojovací podsít s datovým provozem mezi PIM-DM routery byl vybrán jeden z neveřejných IP rozsahů, konkrétně 172.17.5.216/29.

```

172.17.5.217 - PIM DR multicast. zdrojů
172.17.5.218 - PIM router 1. odběratele
172.17.5.219 - PIM router 2. odběratele
172.17.5.221 - Sonda pro VLAN 210

```

obr. 35 - alokace 172.17.5.216/29 na VLAN 210

3.5.2 IGMP doména

Představuje přístupovou část sítě, ve které IPTV set top boxy prostřednictvím IGMP dotazů komunikují přímo s hraničním (PIM-DM) směrovačem poskytovatele. Komunikace probíhá ve vyhrazených IPTV VLAN (VLAN 155x), každá s vlastní podsítí. Ve zjednodušené topologii uvedené výše jsou znázorněny dvě tyto podsítě, ale může jich být i více. Důvodem rozdělení je snaha usnadnit logickou orientaci v přidělovaných IP rozsazích, dále také zmenšení broadcastové domény.

Protože musí tyto podsítě pojmout všechny klientské set top boxy, je zapotřebí zvolit širší IP rozsah než v případě propojovací podsítě pro PIM doménu. V tomto případě byl zvolen rozsah 172.17.240.0/23. Adresy jsou přitom klientským zařízením přidělovány DHCP serverem, který není ve správě poskytovatele. DHCP požadavky jsou výchozí branou – tedy hraničním směrovače – přesměrovávány na DHCP server dodavatele IPTV platformy.

```

172.17.240.1          PIM-DM router poskytovatele / výchozí brána subnetu
172.17.240.2 až .8   rezerva
172.17.240.9        sonda pro IGMP doménu
172.17.240.10       dynamický IP rozsah pro set top boxy (DHCP)
až
172.17.241.254      (celkem až 500 uzlů)

```

obr. 36 - alokace 172.17.240.0/23 na VLAN 1559

PIM router poskytovatele tedy plní funkci *IGMP Querier* - v periodických intervalech odesílá do připojených podsítí *IGMP Query* a podle odpovědí *IGMP Report/Leave* zákaznických IPTV set top boxů udržuje stav tabulky, podle které buduje potřebné distribuční stromy.

Přestože podpora *IGMP snooping*u není nutnou podmínkou pro fungování multicastu, v praxi se switche bez této funkcionality neobejdou, viz. popis mechanismu v jedné z předchozích kapitol.

Výpis IGMP snoopingu z přístupového switche ukazuje, že je aktivní na IPTV VLAN 1559, a to v druhé verzi protokolu. Funkce *Querier* je vypnuta, switche tedy pouze přeposílá *IGMP Queries* hlavního routeru.

```

Vty-0#show ip igmp snooping
Service Status:           Enabled
Querier Status:           Disabled
Leave proxy status:       Disabled
Priority:                  5
Query Count:              2
Query Interval:          125 sec
Query Max Response Time: 10 sec
Router Port Expire Time: 300 sec
Immediate Leave Processing: Disabled
IGMP Snooping Version:   Version 2
-----
VLAN 1559:
-----
IGMP snooping: Enabled

```

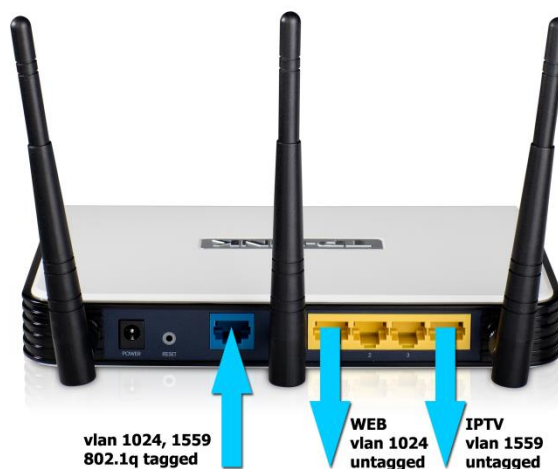
obr. 37 - výpis IGMP snoopingu

CPE

Multicastové a unicastové rámce putují po síti odděleny v různých VLAN. Před použitím v klientských zařízeních je proto nutné je zbavit 802.1q tagů, jelikož jen nízké procento z těchto zařízení s nimi dokáže pracovat.

Rozdělení příchozích VLAN do různých výstupních portů se speciálním využitím a potřebné odstranění 802.1q VLAN tagů (tzv. nativní provoz) obstarává tzv. CPE (*Customer-premises equipment*) jednotka. Možností realizace je více, v popisovaném případě byl

jako CPE jednotka použit domácí router *TP-Link WR941ND* s alternativním firmwre na bázi Linuxu - *OpenWRT*. Všechny porty zařízení pracují v módu bridge, směrovač tedy supluje funkci inteligentního switche. Díky linuxovému jádru je možné tuto CPE jednotku využít i na analýzu datového provozu IPTV boxu, který CPE jednotkou prochází.



obr. 38 - CPE jednotka

Konfigurace CPE jednotky viz. elektronická příloha č. 2.

3.5.3 Sondy

Sondy jsou routery se síťovým rozhraním v některých z IPTV VLAN. Umožňují tak logovat průchozí datový provoz a díky tomu lépe analyzovat chování sítě. V obou případech se jedná o linuxové routery na platformě x86. Jeden z nich je umístěn v pražském Sitelu (SONDA1, IP 172.17.5.22 na VLAN 210), největším propojovacím uzlu v České republice, který je zároveň předávacím místem multicastových video streamů. Další je jedním z lokálních routerů internetového poskytovatele v přístupové IGMP doméně (SONDA2, IP 172.17.5.22 na VLAN 1559).

```
root@sonda1:~# ip route
239.250.1.6 dev eth0 scope link
239.250.1.22 dev eth0 scope link
239.250.1.5 dev eth0 scope link
239.250.4.152 dev eth0 scope link
239.250.1.4 dev eth0 scope link
239.250.1.3 dev eth0 scope link
239.250.1.1 dev eth0 scope link
239.250.2.38 dev eth0 scope link
239.250.1.62 dev eth0 scope link
239.250.2.6 dev eth0 scope link
239.250.1.29 dev eth0 scope link
239.250.1.28 dev eth0 scope link
default via 172.17.16.1 dev eth0
```

obr. 39 - výchozí brány pro IPTV

Pro zachytávání provozu byl použit linuxový nástroj *tcpdump*.

3.6 Síťové prvky a jejich konfigurace

3.6.1 PIM-DM doména

Jen malá část z PIM domény je opravdu ve správě poskytovatele internetu. Jedná pouze o hraniční PIM router. Ostatní síťové prvky jsou ve správě poskytovatele technických streamů, a proto nejsou ve zjednodušeném nákresu síťové infrastruktury zobrazeny.

Hraničním síťovým prvkem je switch *HP A5800-24G*. Jedná se výkonné zařízení, které mimo standardních přepínacích funkcí disponuje i pokročilou funkcionalitou na síťové i transportní vrstvě, která mu dovoluje směřovat veškerý provoz směrovacími



obr. 41 - HP 5800-24G

protokoly PIM, RIP, OSPF, BGP nebo vystupovat jako *IGMP Querier*. Konfigurace tohoto klíčového prvku sítě viz. elektronická příloha č. 2, datasheet s podrobnou technickou specifikací je přiložen jako příloha č. 3. Tabulky s nastavením zde pro přehlednost uvádějí pouze příkazy bezprostředně se týkající typu zařízení a vícesměrového vysílání.

V konfiguraci síťového prvku je potřeba nejdříve nakonfigurovat samotná síťová rozhraní.

Multicastový provoz přichází přes VLAN interface 210 a je směřován do zákaznických VLAN 1553 a 1559. Na všech klientských IPTV VLAN je zapnuto zpracování IGMP. Veškeré DHCP požadavky z IPTV VLAN (tedy požadavky set top boxů) jsou přeposílány na

```
<HP>display version
HP Comware Platform Software
Comware Software, Version 5.20, Release 1211P01
Copyright 2010-2011 Hewlett-Packard Devel. co.
HP A5800-24G Switch uptime is ###

HP A5800-24G Switch with 2 Processors
1024M bytes SDRAM
4M bytes Nor Flash Memory
512M bytes Nand Flash Memory
Config Register points to Nand Flash

Hardware Version is Ver.B
CPLD Version is 003
```

obr. 40 - display version PIM routeru

server poskytovatele technických streamů, který jejich adresování spravuje vzdáleně. PIM probe-interval byl kvůli zvýšení stability zkrácen ze standardních 5 sekund na 1 sekundu. Odeslané PIM Hello pakety inzerují prioritu routeru jako 2. Pro zvýšení stability byla také prodloužena životnost zdroje. Následuje statické směrování multicastových zdrojů na PIM router poskytovatele streamů a úprava QoS, a to přemapováním hodnoty DSCP v odchozích IPTV provozu.

```
#
vlan 210
  name multicast-zdroj
#
vlan 1553
  name iptv-vlan-1
#
vlan 1559
  name iptv-vlan-2
#
interface Vlan-interface210
  ip address 172.17.5.219 255.255.255.248
  undo pim state-refresh-capable
  pim dm
#
interface Vlan-interface1553
  ip address 172.17.228.1 255.255.254.0
  dhcp select relay
  dhcp relay server-select 1
  igmp enable
#
interface Vlan-interface1559
  ip address 172.17.240.1 255.255.254.0
  dhcp select relay
  dhcp relay server-select 1
  igmp enable
#
  dhcp relay server-group 1 ip 172.17.16.15
#
pim
  probe-interval 1
  hello-option dr-priority 2
  hello-option neighbor-tracking
  source-lifetime 10000
#
  ip route-static 172.16.0.0 255.240.0.0 172.17.5.216
  ip route-static 172.17.2.0 255.255.255.0 172.17.5.217
  ip route-static 172.17.4.0 255.255.255.0 172.17.5.217
  ip route-static 172.17.8.0 255.255.255.0 172.17.5.217
  ip route-static 172.17.16.0 255.255.255.0 172.17.5.217
#
traffic behavior iptvprio
  remark dot1p 5
  remark dscp 63
#
qos policy iptvprio
  classifier any behavior iptvprio
#
qos vlan-policy iptvprio vlan 1553 outbound
qos vlan-policy iptvprio vlan 1559 outbound
```

obr. 42 - konfigurace hraničního PIM routeru poskytovatele

3.6.2 IGMP doména

IGMP doména je již plně v rukou poskytovatele. Multicastové datagramy v IPTV VLAN už nejsou dále směrovány, VLAN je zakončena až na koncových zařízeních (klientský set top box), provoz je tedy zcela transparentní na linkové vrstvě.

IPTV boxy komunikují se svým *PIM Designated Routerem* prostřednictvím protokolu IGMP ve verzi 2. K zabránění nekontrolovaného šíření multicastových streamů je proto nutné, aby všechna zařízení po cestě uměla s IGMP pracovat, tedy podporovala techniku *IGMP Snooping*. Konkrétně se jedná o soustavu agregačních, ale především přístupových přepínačů.

Edge-core ES3528M je typickým představitelem spravovatelného L2 switchu s omezenou znalostí transportní vrstvy, která mu umožňuje efektivně nakládat i s multicastovými datagramy (funkce *IGMP snooping*).

Tyto síťové prvky tvoří páteř IGMP

domény a jsou na síti nasazeny obvykle ve stovkách kusů. Konfigurace přístupového switchu viz. elektronická příloha č. 2, datasheet s podrobnou technickou specifikací je přiložen jako příloha č. 3. Tabulky s nastavením pro přehlednost uvádějí pouze příkazy bezprostředně se týkající typu zařízení a vícesměrového vysílání.

Na přepínači je vytvořeno několik různých VLAN, konkrétně VLAN 1552 (multicastový provoz), VLAN 1652 (management VLAN pro CPE jednotky), dále VLAN 1502 (management přepínačů) a konečně VLAN 2047 (klientská internetová VLAN). Důsledné dělení do různých VLAN nejenže zmenšuje broadcastovou doménu, ale také zvyšuje přehlednost celého řešení přenosu různých služeb a typů síťového provozu na sdílené

telekomunikační síti. Všechny rámce přicházejí jako 802.1q *tagged* na páteřním portu 1/28,



obr. 43 – přístupový switch Edge-core es3528m

```
Vty-0#show version
Unit 1
Serial Number:           A836040906
Hardware Version:        R01
Chip Device ID:          Marvell 98DX106-B0
                          88E6095[F]
EPLD Version:            0.07
Number of Ports:         28
Main Power Status:       Up
Redundant Power Status:  Not present

Agent (Master)
Unit ID:                  1
Loader Version:           1.0.0.2
Boot ROM Version:         1.0.0.8
Operation Code Version:   1.4.20.1
```

obr. 44 - show version access switchu

```

!
queue mode strict
!
vlan database
VLAN 1 name DefaultVlan media ethernet state active
VLAN 1502 name management media ethernet state active
VLAN 1552 name iptv media ethernet state active
VLAN 1652 name cpe media ethernet state active
VLAN 2047 name klient media ethernet state active
!
interface ethernet 1/1
description odvod-klient
switchport allowed vlan add 2047 untagged
switchport native vlan 2047
switchport allowed vlan remove 1
switchport allowed vlan add 1552,1652 tagged
!
interface ethernet 1/28
description privod_optika
switchport acceptable-frame-types tagged
switchport allowed vlan add 1502,1552,1652,2047 tagged
!
ip igmp snooping priority 5
no ip igmp snooping vlan 1
no ip igmp snooping vlan 1502
no ip igmp snooping vlan 1652
no ip igmp snooping vlan 2047
!
no map ip precedence
map ip dscp

```

obr. 45 - multicast konfigurace access switche

klientská VLAN 2047 se zbavuje tagu (*untagged*) na všech klientských portech 1/1 až 1/24 (kvůli zachování zpětné kompatibility) a naopak. Na klientských portech jsou dále přítomny tagované IPTV a CPE management VLAN pro CPE jednotku. Následuje zvýšení priority pro IPTV VLAN (proprietární pro Edge-core) a zákaz provádění *IGMP snooping* na jiné než IPTV VLAN, protože jinde než tam by to bylo zbytečné. Následuje nastavení mapování DSCP hodnoty z IP paketu do CoS pole zpracovaného rámce, viz. kapitola pojednávající o QoS.

```

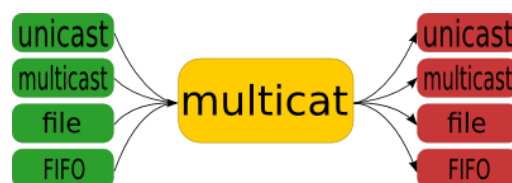
Vty-0#show ip igmp snooping
Service Status:           Enabled
Querier Status:           Disabled
Leave proxy status:        Disabled
Priority:                   5
Query Count:               2
Query Interval:            125 sec
Query Max Response Time:  10 sec
Router Port Expire Time:  300 sec
IGMP Snooping Version:    Version 2
-----
VLAN 1552:
-----
IGMP snooping: Enabled

```

obr. 46 - výpis show ip igmp snooping

3.6.3 Sondy

Jak už bylo řečeno výše, sondy jsou linuxové stanice s distribucí Debian, jejichž síťová rozhraní ležící na zkoumané podsíti. Prostřednictvím nástrojů *tcpdump* a *multicat* (multicastová alternativa k nástroji *Netcat* -



obr. 47 - multicat

dovoluje mimo jiné naslouchat na multicastové skupině a odebírat tak jí adresované UDP datagramy) je tak možné zkoumat nejen broadcastový provoz na dané podsíti, ale také simulovat činnost koncového zařízení, a tedy přijímat a analyzovat příchozí multicastové datagramy. Samotné použití je poměrně triviální.

Protože na sondě neběží žádný multicastový směrovací démon, je potřeba staticky směrovat na všechny odebírané skupiny.

```
sonda2:~# ip route add 239.250.1.1 via eth0
sonda2:~#
sonda2:~# ip route
sonda2:~# 239.250.1.1 dev eth0 scope link
```

obr. 49 - statické nasměrování skupinové adresy

Multicat má, stejně jako Netcat, na poli analýzy TCP/IP široké využití. Primárně se jedná o nástroj, který dokáže číst a zapisovat z/do UDP socketu. Navíc ale také umožňuje přijaté pakety zapsat do souboru a naopak, popř. datagram přeadresovat na unicast.

```
sonda2:~# multicat @239.250.1.1:1234 /dev/null
sonda2:~# multicat
sonda2:~# multicat @239.250.1.1:1234 stream.ts
sonda2:~# multicat
sonda2:~# multicat @239.250.1.1:1234 10.0.0.1:1001
```

obr. 48 - použití nástroje multicat

V popisované aplikaci se využívá především jeho schopnost zachytit multicastový datagram.

```
root@sonda2:~# tcpdump -ni eth0 host 239.250.1.1
tcpdump: verbose output suppressed, use -v or -vv
for full protocol decode
listening on eth0, link-type EN10MB (Ethernet),
capture size 96 bytes
00:44:19.360852 IP 172.17.4.2.37266 >
239.250.1.1.1234: UDP, length 1316
00:44:19.362570 IP 172.17.4.2.37266 >
239.250.1.1.1234: UDP, length 1316
00:44:19.364279 IP 172.17.4.2.37266 >
239.250.1.1.1234: UDP, length 1316
3 packets captured
3 packets received by filter
0 packets dropped by kernel
```

obr. 50 – tcpdump odebírané skupiny

Protože zároveň zobrazuje jeho sekvenční číslo id (ve

verbose módu), z posloupnosti příchozích datagramů je možné získávat ztrátovost přenosu.

3.7 Analýza provozu

3.7.1 PIM-DM doména

Jak už bylo uvedeno v jedné z předchozích kapitol, multicastové datagramy jsou na PIM doméně routovány protokolem PIM v režimu *Dense mode*. Ten najde uplatnění všude tam, kde nejsou multicastové zdroje a jejich příjemci geograficky příliš vzdáleni, existuje pouze několik zdrojů, zato mnoho příjemců a jsou přenášeny velké datové objemy. A právě proto je PIM-DM pro IPTV aplikace velmi vhodný a často využívaný, přestože v prostředí internetu je jeho využití spíše na ústupu ve prospěch alternativního Sparse módu, který ve velkém měřítku přeci jen lépe škáluje.

Hello messages

Zprávy Hello PIM-DM obecně využívá k nalezení ostatních PIM routerů. *Hold Time Option* udává dobu, po kterou příjemce této zprávy nesmí odesílatele vyřadit ze seznamu sousedů, resp. jak dlouho budou informace obsažené v *Hello messages* platné. *Hold time* bývá obvykle 3,5ti násobkem *Hello period*, tedy intervalem zasílání Hello zpráv. V popisovaném nastavení je to opravdu 1 minuta a 45 sekund, tedy 105 sekund, což je přesně trojnásobek *Hello period* nastavené na 30 sekund.

LAN Prune Delay Option je možné využít k úpravě zpoždění propagace J/P zpráv na lokální síti. Zatímco hodnotu T-bitu PIM-DM nevyužívá, hodnota *Override interval* musí být nutně vyšší než hodnota *LAN delay* (administrátorem konfigurovaná maximální předpokládaná doba šíření zprávy) tak, aby soused downstream routeru požadujícího *Prune* stihl zareagovat a požádat o případný *Join* ještě před přerušáním preposílání datagramů na společnou síť.

Generation ID Option je náhodná hodnota vygenerovaná pro rozhraní, na které je *Hello message* zaslána.

```
sonda1:~# tcpdump -ni eth1.210 host 172.17.5.217 | grep Hello
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth1.210, link-type EN10MB (Ethernet), capture size 65535 bytes
23:44:05.821432 IP 172.17.5.217 > 224.0.0.13: PIMv2, Hello, length 42
23:44:35.821972 IP 172.17.5.217 > 224.0.0.13: PIMv2, Hello, length 42
```

obr. 51 - periodické zasílání *Hello messages*

Zprávy jsou zasílány PIM routery na skupinovou adresu všech ostatních routerů na lokální síti a opakují se periodicky každých 30 s.

```

sondal:~# tcpdump -vni eth1.210 | grep -A 9 -B 2 Hello
tcpdump: listening on eth1.210, link-type EN10MB (Ethernet)
23:35:35.804352 IP (tos 0xc0, ttl 1, id 54491, offset 0, flags [none], proto PIM
(103), length 62)
  172.17.5.217 > 224.0.0.13: PIMv2, length 42
    Hello, cksum 0x8d14 (correct)
      Hold Time Option (1), length 2, Value: 1m45s
      LAN Prune Delay Option (2), length 4, Value:
        T-bit=0, LAN delay 500ms, Override interval 2500ms
      DR Priority Option (19), length 4, Value: 1
      Generation ID Option (20), length 4, Value: 0x7185d3b6
      State Refresh Capability Option (21), length 4, Value: v1, interval 1m
    --
23:35:37.183058 IP (tos 0xc0, ttl 1, id 2774, offset 0, flags [none], proto PIM
(103), length 54)
  172.17.5.219 > 224.0.0.13: PIMv2, length 34
    Hello, cksum 0xa956 (correct)
      Hold Time Option (1), length 2, Value: 1m45s
      LAN Prune Delay Option (2), length 4, Value:
        T-bit=1, LAN delay 500ms, Override interval 2500ms
      DR Priority Option (19), length 4, Value: 2
      Generation ID Option (20), length 4, Value: 0xd712d33a
    --
23:35:53.795895 IP (tos 0xc0, ttl 1, id 17127, offset 0, flags [none], proto PIM
(103), length 54)
  172.17.5.218 > 224.0.0.13: PIMv2, length 34
    Hello, cksum 0xb6bf (correct)
      Hold Time Option (1), length 2, Value: 6m7s
      LAN Prune Delay Option (2), length 4, Value:
        T-bit=0, LAN delay 500ms, Override interval 2500ms
      DR Priority Option (19), length 4, Value: 1
      Generation ID Option (20), length 4, Value: 0x35f4e5eb

```

obr. 52 – zachycené *Hello* zprávy PIM routerů na společné doméně

State Refresh Capability Option udává interval odesílání *State refresh* zpráv, které jsou odesílány PIM routery směrem ke kořenu distribučního stromu a slouží k prodloužení *Prune state* nadřazeného směrovače tak, aby nedocházelo k periodické záplavě, která je typická pro PIM-DM.

Join/Prune/Graft zprávy

PIM zprávy typu *Prune* jsou v případě PIM-DM používány k odmítnutí záplavy multicastových streamů skupiny G, zdroje S. Tyto zprávy jsou odesílány zpět upstream sousedovi, tedy na RPF (*Reverse Path Forwarding*) rozhraní PIM routeru.

```

sondal:~# tcpdump -nvi eth1.210 net 224.0.0.0/4 | grep -B 6 -A 1 239.250.1.8
tcpdump: listening on eth1.210, link-type EN10MB (Ethernet), capture size 65535
bytes

01:34:49.626013 IP (tos 0xc0, ttl 1, id 57851, offset 0, flags [none], proto PIM
(103), length 54)
  172.17.5.218 > 224.0.0.13: PIMv2, length 34
    Join / Prune, cksum 0x8590 (correct), upstream-neighbor: 172.17.5.217
      1 group(s), holdtime: 5m
        group #1: 239.250.1.8, joined sources: 0, pruned sources: 1
          pruned source #1: 172.17.4.2

01:34:51.570747 IP (tos 0xc0, ttl 1, id 48273, offset 0, flags [none], proto PIM
(103), length 54)
  172.17.5.219 > 224.0.0.13: PIMv2, length 34
    Join / Prune, cksum 0x86bc (correct), upstream-neighbor: 172.17.5.217
      1 group(s), holdtime: 0s
        group #1: 239.250.1.8, joined sources: 1, pruned sources: 0
          joined source #1: 172.17.4.2

```

obr. 53 - použití Join/Prune zpráv v PIM-DM

V případě, že více PIM routerů sdílí společný LAN segment, dochází často k následující situaci:

1. PIM router č.1 (172.17.5.218) žádá upstream router (172.17.5.217) o ukončení zasílání multicastových dat (Prune zpráva) pro skupinu G (239.250.1.8), protože již pro ni sám nemá žádné odběratele
2. protože je tato zpráva adresována všem směrovačům na segmentu (224.0.0.13), zachytí ji i ostatní PIM routery
3. PIM router č.2 (172.17.5.219) má stále alespoň jednoho aktivního odběratele skupiny G, proto musí Prune zprávu PIM routeru č.1 „přebít“ Join zprávou, která zajistí nepřerušování datového toku.

V žádném jiném případě PIM protokol v Dense módu zprávy Join nepoužívá.

Dalším typem zpráv protokolu PIM jsou tzv. *Graft messages*, kterými PIM router vyjadřuje zájem o opětovné přihlášení k již dříve odhlášené skupině.

```

<HP>display pim control-message counters
VPN-Instance: public net
PIM global control-message counters:
      Received      Sent      Invalid
Register      0          0          0
Register-Stop  0          0          0
Probe         0          0          0

PIM control-message counters for interface: Vlan-interface210
      Received      Sent      Invalid
Assert      0          0          0
Graft      0          158489    0
Graft-Ack   146864    0          0
Hello     676416    349601    137533
Join/Prune 13253891 6875204    1566783
State-Refresh 4          0          4
BSR         56057    0          0
C-RP        0          0          0

```

obr. 54 - statistika Join/Prune/Graft zprávna hraničním routeru

Zprávy Graft jsou PIM routerem zasílány na rozhraní upstream souseda jako unicast, není možné je tedy snadno odchytil, PIM router poskytovatele navíc neumožňuje Graft zprávy cíleně analyzovat. Ze statistiky odeslaných zpráv je ale zřejmé, že k jejich generování dochází v průměru dvakrát méně často než v případě PIM Hello zpráv. Protože jsou Hello zprávy generovány periodicky každých 30 sekund, lze říci, že v popisovaném případě se PIM router poskytovatele přihlašuje k dříve odhlášenému odběru skupiny G v průměru jedenkrát za minutu.

Z přehledu zpráv je dále patrné, že z PIM routeru poskytovatele opravdu neodchází žádné *State-refresh messages*, které mají za úkol prodloužit *Prune state* upstream routeru pro danou skupinu G. Tato optimalizační technika, která PIM-DM přibližuje k PIM-SM není pro aplikace IPTV vhodná, jelikož oddaluje v tomto případě vítanou záplavu multicastovými streamy.

Směrování multicastových datagramů

Kompletní směrovací tabulka na hraničním PIM routeru vytvořená protokolem PIM-DM čítá celkem 224 záznamů notace (S,G), po jednom záznamu pro každý televizní kanál a rádiovou stanicí.

Po vyfiltrování konkrétní skupinové adresy (která představuje video stream ČT1) je zobrazen jediný (S,G) záznam, který odkazuje na vybudovaný distribuční strom se specifickým zdrojem 172.17.4.2. *Upstream neighbor* (nejbližší router směrem ke kořenu distribučního stromu) i

RPF prime neighbor (router, na který odkazuje porovnání záznamů s unicastovou směrovací tabulkou) jsou jeden a tentýž PIM router 172.17.5.217, v topologii tedy nedochází ke smyčce. Značka ACT říká, že daný záznam popisuje skutečný datový provoz. Z pohledu směrovacího protokolu PIM

neexistuje žádný downstream interface, tedy odchozí síťové rozhraní k dalšímu PIM routeru. Všechny multicastové datagramy jsou totiž předávány koncovým uzlům přes protokol IGMP.

Záznam z obecné směrovací tabulky zobrazuje i odchozí síťová IGMP rozhraní, tedy uživatelské IPTV VLAN 1553 a VLAN 1559.

```
<HP>display pim routing-table 239.250.1.1
Total 224 (S, G) entries
Total matched 1 (S, G) entry

(172.17.4.2, 239.250.1.1)
  Protocol: pim-dm, Flag: EXT ACT
  UpTime: 4w:5d
  Upstream interface: Vlan-interface210
    Upstream neighbor: 172.17.5.217
    RPF prime neighbor: 172.17.5.217
  Downstream interface(s) information: None
```

obr. 55 - záznam skupiny v PIM routovací tabulce

```
<HP>display multicast routing-table 239.250.1.1
Multicast routing table of VPN-Instance: public net
Total 221 entries ,1 matched

00001. (172.17.4.2, 239.250.1.1)
  Uptime: 2w:0d
  Upstream Interface: Vlan-interface210
  List of 2 downstream interfaces
    1: Vlan-interface1553
    2: Vlan-interface1559
```

obr. 56 - záznam skupiny v obecné routovací tabulce

Mimo multicastových skupin reprezentujících jednotlivé televizní a rádiové stanice set top boxy naslouchají také na skupinových adresách dedikovaných pro tzv. *SAP Announcements*, viz. dále.

```
<HP>display multicast routing-table 224.2.2.8
Multicast routing table of VPN-Instance: public net
Total 224 entries ,1 matched

00001. (172.17.8.86, 224.2.2.8)
  Uptime: 7w:4d
  Upstream Interface: Vlan-interface210
  List of 2 downstream interfaces
    1: Vlan-interface1553
    2: Vlan-interface1559
```

obr. 57 - záznam SAP skupiny v obecné routovací tabulce

3.7.2 IGMP doména

Jak už bylo uvedeno výše, IPTV datové přenosy putují k zákazníkovi speciálně vyčleněnými klientskými IPTV podsítěmi. Ve skutečnosti je jich celkem devět, každá z nich přitom pokrývá jednu geografickou oblast, může ale být samozřejmě i pouze jediná – záleží na rozhodnutí poskytovatele. Ve zjednodušené topologii jsou zakresleny dvě z nich, IPTV VLAN 1553 a IPTV VLAN 1559. Každá z nich propojuje koncová uživatelská zařízení (CPE jednotky, resp. set top boxy Motorola za nimi) s VLAN síťovým rozhraním na hraničním PIM routeru poskytovatele – multicastový IPTV provoz tedy v síti poskytovatele již dále není směrován. Veškerý multicastový provoz na těchto podsítích je řízen protokolem IGMPv2.

```
<HP>display igmp interface Vlan-interface 1553 verbose
Interface information of VPN-Instance: public net

Vlan-interface1553(172.17.228.1):
  IGMP is enabled
  Current IGMP version is 2
  Value of query interval for IGMP(in seconds): 60
  Value of other querier present interval for IGMP(in seconds): 125
  Value of maximum query response time for IGMP(in seconds): 10
  Value of last member query interval(in seconds): 1
  Value of startup query interval(in seconds): 15
  Value of startup query count: 2
  General query timer expiry (hours:minutes:seconds): 00:00:28
  Querier for IGMP: 172.17.228.1 (this router)
  IGMP activity: 60059 joins, 60048 leaves
  Multicast routing on this interface: enabled
  Robustness: 2
  Require-router-alert: disabled
  Fast-leave: disabled
  Ssm-mapping: disabled
  Startup-query-timer-expiry: off
  Other-querier-present-timer-expiry: off
  Proxying interface: None
  Total 11 IGMP Groups reported
```

obr. 58 - nastavení IGMP parametrů klientské IPTV vlan

Dotaz IGMP Query tedy opouští router každých 60 sekund (hodnota *query-interval*). Maximální čas na odpověď je nastavena na 10 sekund (*query-response-interval*). Tato hodnota je zároveň uložena v hlavičce samotného Query dotazu. *General query timer expiry* ukazuje čas zbývající k opětovnému odeslání IGMP Query. Hodnota *Robustness* dále určuje citlivost systému na ztrátu paketů. IGMP se dokáže zotavit, pokud hodnota *Robustness* zůstane 1 a více, přičemž každý ztracený paket tuto hodnotu sníží o jedna. Pod daným síťovým rozhraním router eviduje celkem 11 různých IGMP skupin. Odběratelů může být samozřejmě více.

Rozložení jedné z nejodebíranějších multicastových skupin 239.250.1.1 (ČT1) ukazuje, že její *uptime* je na jedné IPTV podsíti tři a půl dne, na druhé dokonce několik měsíců – na obecný IGMP Query dotaz routeru vždy dorazí IGMP Report klientského set top boxu pro danou multicastovou skupinu (to také ukazuje na skutečnost, že zákazníci své set top boxy často ani nevypínají). *Last reporter* ukazuje na poslední uzel, od kterého router obdržel *IGMP Report* pro danou skupinu.

```
<HP>display igmp group 239.250.1.1 verbose
Vlan-interface1559(172.17.240.1):
  Total 5 IGMP Groups reported
  Group: 239.250.1.1
  Uptime: 3d:14h
  Expires: 00:01:54
  Last reporter: 172.17.240.9
  Last-member-query-counter: 0
  Last-member-query-timer-expiry: off
  Version1-host-present-timer-expiry: off
Vlan-interface1553(172.17.228.1):
  Total 10 IGMP Groups reported
  Group: 239.250.1.1
  Uptime: 11w:0d
  Expires: 00:01:53
  Last reporter: 172.17.228.10
  Last-member-query-counter: 0
  Last-member-query-timer-expiry: off
  Version1-host-present-timer-expiry: off
```

obr. 59 - statistika IGMP odběratelů multicastové skupiny

Každých 60 sekund tedy router odesílá obecný *IGMP Query* dotaz na všechny své *IGMP Enabled* virtuální síťová rozhraní. Tyto dotazy jsou odesílány na multicastovou adresu 224.0.0.1, tedy „všechny uzly na místní síti“. Query požadavek tedy v jeden okamžik obdrží všechny aktivní set top boxy na síti. Kvůli možnému zahlcení sítě ale neodpovídají okamžitě, pouze spustí interní časovač s náhodným zpožděním, maximálně však s hodnotou *query-response-interval* definovanou routerem v příchozím Query požadavku tak, aby náhodou nedošlo k přerušení přeposílání dat skupiny. V mezičase poslouchá odpovědi ostatních odběratelů. Pokud zaregistruje *IGMP Report* pro některou ze svých skupin, svou zpožděnou odpověď už neodešle vůbec.

```
root@sonda2:~# tcpdump -nvvvi br-iptv ip proto 2
tcpdump: listening on br-iptv, link-type EN10MB (Ethernet), capture size 65535
bytes
00:57:38.854797 IP (tos 0xfc, ttl 1, id 52820, offset 0, flags [none], proto
IGMP (2), length 32, options (RA))
  172.17.240.1 > 224.0.0.1: igmp query v2
00:57:39.661420 IP (tos 0xc0, ttl 1, id 0, offset 0, flags [DF], proto IGMP (2),
length 32, options (RA))
  172.17.241.251 > 224.2.2.8: igmp v2 report 224.2.2.8
00:57:47.365884 IP (tos 0xc0, ttl 1, id 0, offset 0, flags [DF], proto IGMP (2),
length 32, options (RA))
  172.17.241.251 > 239.250.2.24: igmp v2 report 239.250.2.24
3 packets captured
12 packets received by filter
```

obr. 60 - zachycení IGMP komunikace set top boxu

Po startu set top boxu Motorola zařízení nejdříve požádá vzdálené servery o přidělení IP adresy skrze klasický *DHCP Request*. Tyto servery nejsou ve správě poskytovatele, ale společnosti Nangu.TV, která provoz IPTV platformy zajišťuje.

```

root@sonda2:~# tcpdump -nvi br-iptv port 67 or port 68
tcpdump: listening on br-iptv, link-type EN10MB (Ethernet), capture size 65535
bytes
17:47:15.032881 IP (tos 0x0, ttl 64, id 0, offset 0, flags [DF], proto UDP (17),
length 576)
    0.0.0.0.68 > 255.255.255.255.67: BOOTP/DHCP, Request from 00:02:9b:93:7f:71,
length 548, xid 0x59a89b15, Flags [none]
    Client-Ethernet-Address 00:02:9b:93:7f:71
    Vendor-rfc1048 Extensions
        Magic Cookie 0x63825363
        DHCP-Message Option 53, length 1: Discover
        MSZ Option 57, length 2: 1500
        Parameter-Request Option 55, length 13:
            Subnet-Mask, Default-Gateway, Domain-Name-Server, Hostname
            Domain-Name, SS, RP, MTU
            BR, YD, NTP, Vendor-Option
            Option 234
        Vendor-Class Option 60, length 16: "Motorola_VIP1003"
17:47:16.006032 IP (tos 0x0, ttl 255, id 31558, offset 0, flags [none], proto
UDP (17), length 369)
    172.17.240.1.67 > 172.17.241.115.68: BOOTP/DHCP, Reply, length 341, hops 1,
xid 0x59a89b15, Flags [none]
    Your-IP 172.17.241.115
    Gateway-IP 172.17.240.1
    Client-Ethernet-Address 00:02:9b:93:7f:71
    Vendor-rfc1048 Extensions
        Magic Cookie 0x63825363
        DHCP-Message Option 53, length 1: Offer
        Server-ID Option 54, length 4: 172.17.16.15
        Lease-Time Option 51, length 4: 21600
        Subnet-Mask Option 1, length 4: 255.255.254.0
        Default-Gateway Option 3, length 4: 172.17.240.1
        Domain-Name-Server Option 6, length 4: 172.17.16.15
        Domain-Name Option 15, length 21: "cho01.iptv.grapesc.cz"
        Vendor-Option Option 43, length 42:
1.1.49.2.1.49.3.17.49.55.50.46.49.55.46.56.46.50.53.58.49.57.57.57.55.10.15.50.5
0.52.46.50.46.50.46.56.58.50.50.50.50.50

```

obr. 61 - přidělení IP set top boxu službou DHCP vzdáleným serverem

Veškeré DHCP požadavky ze zákaznických IPTV podsítí jsou přitom přeposílány (*DHCP Relay*) na DHCP server poskytovatele IPTV platformy (172.17.16.15). Součástí *DHCP Reply* je kromě klasických atributů také speciální *Option 43*, díky které set top box získává informaci o aktuálně dostupných SW aktualizacích. Po porovnání řetězce s řetězcem uloženým ve své paměti může přikročit ke stáhnutí nového software.

Set top box bývá zpravidla připojen ke dvěma multicastovým skupinám. První je adresa skupiny z rozsahu *Administratively Scoped addresses*, tedy privátního rozsahu, na kterém jsou adresovány jednotlivé televizní a rádiové stanice.

```

root@sonda2:~# tcpdump -nvi br-iptv host 239.250.2.24
tcpdump: WARNING: br-iptv: no IPv4 address assigned
tcpdump: listening on br-iptv, link-type EN10MB (Ethernet), capture size 65535
bytes
00:38:36.969417 IP (tos 0xff,CE, ttl 61, id 61411, offset 0, flags [none], proto
UDP (17), length 1344)
    172.17.4.2.37266 > 239.250.2.24.1234: UDP, length 1316
00:38:36.971771 IP (tos 0xff,CE, ttl 61, id 61412, offset 0, flags [none], proto
UDP (17), length 1344)
^C
2 packets captured
17 packets received by filter
0 packets dropped by kernel

```

obr. 62 - multicastové IPTV datagramy

Příchozí UDP datagramy mají hraničním routerem namapovány vysoké hodnoty *Type of Service*, aby byly síťovými prvky na cestě upřednostňovány před ostatním datovým provozem.

Set top boxy také odebírají data skupiny 224.2.2.8, kterou využívá tzv. SDP - *Session Description Protocol* [23]. Pomocí něj poskytovatel IPTV platformy informuje všechny aktivní set top boxy o parametrech a formátech přenosu, distribuovaných kanálech apod. Samotná obrazová data přitom tímto kanálem nikdy šířena nejsou.

```

root@sonda2:~# tcpdump -nvvi br-iptv host 224.2.2.8
tcpdump: listening on br-iptv, link-type EN10MB (Ethernet), capture size 65535
bytes
00:47:07.736659 IP (tos 0xfc, ttl 7, id 0, offset 0, flags [DF], proto UDP (17),
length 177)
    172.17.8.86.22222 > 224.2.2.8.22222: [udp sum ok] UDP, length 149
00:47:07.741796 IP (tos 0xfc, ttl 7, id 0, offset 0, flags [DF], proto UDP (17),
length 328)
    172.17.8.86.22222 > 224.2.2.8.22222: [udp sum ok] UDP, length 300
^C
2 packets captured
18 packets received by filter
0 packets dropped by kernel

```

obr. 63 - multicastové SDP datagramy

Vzhledem ke zpoždění, které vzniká pomalou propagací IGMP zpráv po síťových prvcích může dojít k situaci, kdy se set top box stane odběratelem i více multicastových IPTV streamů současně vlivem rychlého přepínání programů, na které síť nedokáže dostatečně rychle zareagovat.

```
Vty-0#show ip igmp snooping groups vlan 1559
VLAN IP Addressses  Member Port  Type                UnReply Query Last Reporter
-----
1559 224.2.2.8         Eth 1/ 1 IGMP Snooping      0 172.17.241.251
1559 239.250.1.43      Eth 1/ 1 IGMP Snooping      0 172.17.241.251
1559 239.250.1.47      Eth 1/ 1 IGMP Snooping      0 172.17.241.251
1559 239.250.1.64      Eth 1/ 1 IGMP Snooping      0 172.17.241.251
```

obr. 64 - přehled odebíraných multicastových skupin na přístupovém switchi

Po obdržení zprávy *IGMP Leave* od posledního odběratele multicastové skupiny na příslušné IGMP doméně router poskytovatele (*IGMP Querier*) odesílá specifickou IGMP Query ke zjištění zbývajících odběratelů skupiny na místní síti.

```
root@sonda2:~# tcpdump -nvvvi br-iptv ip proto 2
tcpdump: listening on br-iptv, link-type EN10MB (Ethernet), capture size 65535
bytes

16:23:50.795131 IP (tos 0xc0, ttl 1, id 0, offset 0, flags [DF], proto IGMP (2),
length 32, options (RA))
    172.17.241.251 > 224.0.0.2: igmp leave 239.250.1.59

16:23:50.848422 IP (tos 0xc0, ttl 1, id 28161, offset 0, flags [none], proto
IGMP (2), length 32, options (RA))
    172.17.240.1 > 239.250.1.59: igmp query v2 [max resp time 10] [gaddr
239.250.1.59]

16:23:51.257865 IP (tos 0xc0, ttl 1, id 0, offset 0, flags [DF], proto IGMP (2),
length 32, options (RA))
    172.17.241.251 > 239.250.1.2: igmp v2 report 239.250.1.2

16:23:51.834349 IP (tos 0xc0, ttl 1, id 28183, offset 0, flags [none], proto
IGMP (2), length 32, options (RA))
    172.17.240.1 > 239.250.1.59: igmp query v2 [max resp time 10] [gaddr
239.250.1.59]

4 packets captured
24 packets received by filter
0 packets dropped by kernel
```

obr. 65 – zachycení specific IGMP Query

Zpráva *IGMP Leave* odchází na skupinovou adresu všech routerů na místní síti (224.0.0.1). Zpracují ji ale také všechna síťová zařízení po cestě s funkcí *IGMP Snooping*. V případě posledního odběratele dané skupiny na místní síti *IGMP Querier* (PIM router poskytovatele) reaguje sérií *IGMP Group-Specific Queries* (počet určuje proměnná *Last Member Query Count*) v rozmezí jedné sekundy (určuje proměnná *Last Member Query Interval*). Tyto jsou adresovány konkrétní skupině. V případě, že *IGMP Querier* neobdrží odpověď do jedné sekundy (hodnota *Max Response time* ve stovkách milisekund), přeruší zcela zasílání dat pro danou multicastovou skupinu.

Pro účely obousměrné komunikace a příjem nelineárních služeb komunikuje set top box také prostřednictvím unicastu. Jak už bylo uvedeno výše, IP adresa a všechny ostatní pro komunikaci potřebné atributy jsou mu přiděleny po startu službou DHCP.

```

root@sonda2:~# tcpdump -ni br-iptv ip proto 6
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on br-iptv, link-type EN10MB (Ethernet), capture size 65535 bytes

15:29:30.816978 IP (tos 0xfc, ttl 61, id 37269, offset 0, flags [DF], proto TCP
(6), length 1500)
    172.17.8.137.49261 > 172.17.241.251.23688: Flags [.], cksum 0x69a2
(correct), seq 1449:2897, ack 0, win 5792, options [nop,nop,TS val 1088074605

15:29:30.817119 IP (tos 0xfc, ttl 61, id 37270, offset 0, flags [DF], proto TCP
(6), length 1500)
    172.17.8.137.49261 > 172.17.241.251.23688: Flags [.], cksum 0xe51c
(correct), seq 2897:4345, ack 0, win 5792, options [nop,nop,TS val 1088074605

15:29:30.818283 IP (tos 0x0, ttl 64, id 34834, offset 0, flags [DF], proto TCP
(6), length 52)
    172.17.241.251.23688 > 172.17.8.137.49261: Flags [R], cksum 0xe469
(correct), seq 0, ack 2897, win 28960, options [nop,nop,TS val 60151441 ecr 10

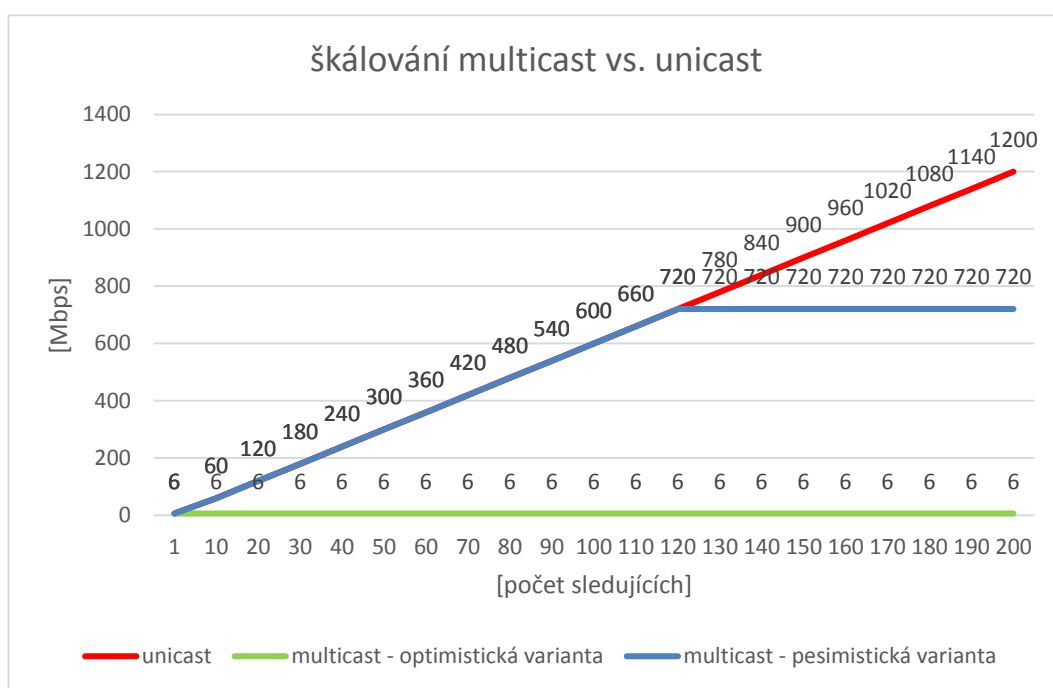
```

obr. 66 - unicastový přenos nelineárního obsahu

Při využití jakékoliv nelineární služby tedy okamžitě dochází k navázání TCP spojení se vzdáleným serverem a přenosu potvrzovaných IP paketů o maximální velikosti MTU 1500 B.

4 MULTICAST VS. UNICAST

Vzhledem k povaze šíření multicastových datagramů po distribučním stromu a jejich replikaci na síťových prvcích nelze bez dodatečné statistické analýzy přesně vyčíslit, jaké zefektivnění přenosu dat vícesměrové vysílání vlastně přináší. Definovat ale můžeme nejlepší a nejhorší případ a následně jej konfrontovat s lineárním škálováním unicastu.



obr. 67 – teoretické škálování multicastu a unicastu

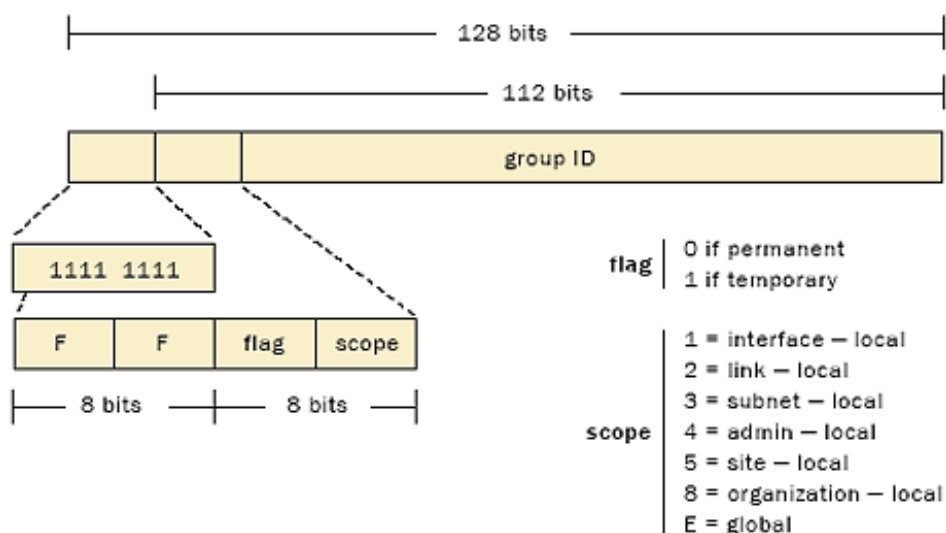
V případě unicastu je situace velmi jednoduchá. Objem datového provozu stoupá lineárně s počtem sledujících. Pokud každý z nich odebírá video stream s datovým tokem 6 Mbit/s, dvě stovky paralelně sledujících generuje provoz 1200 Mbit/s. V případě multicastu nelze snadno rozhodnout, jaký bude výsledný datový tok pro 200 sledujících. Významnou roli hraje uživatelská preference. Pokud by všichni uživatelé sledovali souběžně jediný kanál, celkový datový objem by se rovnal datovému toku tohoto kanálu, tedy 6 Mbit/s. Jednalo by se tedy o nejlepší možnou variantu. Pokud by uživatelé sledovali různé kanály (pesimistická varianta), celkový datový objem by zpočátku kopíroval unicastový průběh až do okamžiku, kdy by každý uživatel odebíral jeden unikátní kanál. V grafu je to 120 sledujících se 120ti televizními kanály a datovým tokem 720 Mbit/s. Při dalším zvyšování počtu sledujících by už ale celkový datový provoz dále nerostl. Reálný průběh by byl zřejmě „někde mezi“ těmito extrémy.

5 MULTICAST A IPV6

Jak už bylo uvedeno v předchozích kapitolách, vícesměrové vysílání hraje v prostředí nového IP protokolu klíčovou roli. Důvodem je naprostá absence všesměrového vysílání, tak jak jej známe z IPv4 a přesunutí jeho role na vysílání vícesměrové. Broadcast se za dlouhá léta rozvoje IPv4 stal noční můrou síťových administrátorů. Byť je ve své podstatě potřebný k navázání síťové komunikace mezi dvěma účastníky, obzvláště na velkých síťových doménách může jeho nekontrolované šíření výrazně ovlivnit chod a výkonnostní parametry celé sítě i síťových prvků, které tuto síť obsluhují. Problémem je jednak zvýšení zátěže CPU, které musí broadcastový provoz vždy zpracovat, dále také vzrůstající počet kolizí domény, která překračuje kolizní doménu a vytváří novou, větší, broadcastovou doménu. Z těchto důvodů tedy není divu, že byl koncept broadcastu v návrhu IPv6 zcela opuštěn.

5.1 Adresování multicastů v IPv6

Multicastová IPv6 adresa se vyznačuje pevným osmibitovým prefixem 1111 1111. Proto je vícesměrové v prostředí IPv6 snadno rozpoznatelné – adresa začíná vždy jako FF. Multicastová adresa navíc nikdy nemůže být použita jako adresa zdrojová.



obr. 68 - IPv6 multicastová adresa [24]

Adresa dále obsahuje pole *Flag*, *Scope* a *Group ID*.

Čtyřbitové pole *Flag*, které označuje různé příznaky dané multicastové IP adresy, najde reálné využití pravděpodobně až v budoucnosti. RFC 4291 [25] zatím definuje jediné použití pole *Flag*, a to rozlišení na permanentně přidělenou adresu organizací IANA (*flag 0 - permanent*) a dočasně přidělenou (*flag 1 - temporary*).

RFC 3306 přidává další příznak signalizující rozšíření IPv6 multicastového adresního schématu o dynamickou alokaci multicastových IPv6 adres.

RFC 3956 dále přináší možnost vtisknutí adresy *Rendezvous Point* přímo do multicastové adresy, což usnadňuje konfiguraci a nasazení multicasu v prostředí IPv6. Není tak nutné konfigurovat adresu RP na každém routeru.

Následující čtyřbitové pole *Scope* určuje tzv. dosah paketu, pro který je takto adresovaný multicastový provoz určen. Slouží jako informace pro směrovače, které tento provoz zpracovávají a doplňuje informace získané od multicastových IPv6 směrovacích protokolů.

hodnota Scope	IPv6 rozsah	ekvivalentní IPv4 rozsah	působnost (Scope)	popis
0	ff00::/16 až ff0f::/16	-	<i>rezervováno</i>	viz. dále
1	ffx1::/16	127.0.0.0/8	Node-local	ekvivalent k <i>loopback interface</i>
2	ffx2::/16	224.0.0.0/24	Link-local	pakety nejsou routovány
3	ffx3::/16	239.255.0.0/16	Subnet-local	přidává podporu směrování
4	ffx4::/16	-	Admin-local	nejmenší jednotka řízené správy
5	ffx5::/16	-	Site-local	<i>deprecated</i>
8	ffx8::/16	239.192.0.0/14	Organization-local	spojení více domén pod jednotnou organizací
E	ffxe::/16	224.0.1.0 až 238.255.255.255	Global	pakety routovány v rámci internetu
F	ffxf::/16	-	<i>rezervováno</i>	-

tab. 3 - přehled adresování IPv6 multicasu

Následuje 112 b definujících ID adresované multicastové skupiny (*Group ID*) v rámci daného rozsahu (*Scope*). Zatímco permanentně přidělená multicastová IPv6 adresa je nezávislá na hodnotě *Scope*, dočasně přidělené adresy existují pouze v tomto daném dosahu.

Např.:

- *FF01::101* reprezentuje všechny NTP servery na stejném interface jako odesílatel (všechny NTP servery na *loopback interface*)
- *FF02::101* reprezentuje všechny NTP servery na stejné linkové vrstvě (broadcastové doméně) jako odesílatel
- *FF0E::101* reprezentuje všechny NTP servery na internetu

Stejné adresy multicastových skupin mohou tedy mimo globální dosah existovat i vícekrát s různými odběrateli.

I v prostředí IPv6 existují tzv. *well-known IPv6 multicast addresses*, tedy adresy staticky přidělené *Internet Assigned Numbers Authority* (IANA) pro použití na lokální síti.

adresa	popis
FF02::1	všichni účastníci na lokální síti (ekvivalent broadcastu)
FF02::2	všechny routery na lokální síti
FF02::5	Shortest Path First routery směrovacího protokolu OSPF
FF02::6	Designated routery směrovacího protokolu OSPF
FF02::9	RIP routery
FF02::D	PIM routery
FF02::12	VRRP agent
FF02::1:2	DHCP servery a relay agenti (RFC 3315)
FF02::101	Network Time Protocol (NTP)

tab. 4 - *well-known IPv6 multicast addresses*

5.1.1 Mapování IPv6 adresy na MAC adresu

Přestože princip statického mapování IP adresy na MAC adresu zůstává v případě IPv6 zachován, konkrétní postup přenosu se od toho použitého v případě IPv4 mírně liší. Vzhledem k tomu, že IPv6 adresa je mnohem delší než fyzická MAC adresa (128 bitů vs. 48 bitů), je jasné, že při mapování bude velká část IPv6 adresy ztracena. Konkrétně se tedy mapuje pouze posledních 32 bitů IPv6 adresy, a to za pevný 16bitový prefix 3333.



obr. 69 - mapování IPv6 multicast adresy na MAC adresu [27]

Např. multicastová IPv6 adresa **FF02::D** (všechny PIM routery na lokální síti) by byla namapována na MAC adresu **33:33:0:0:0:D**.

Vzhledem k velkému nepoměru délky obou adres i v případě IPv6 platí, že může dojít k namapování více IPv6 adres na jedinou MAC adresu. V takovém případě tedy může docházet i k doručování datagramů příjemcům, kteří o něj ani nemají zájem. K rozdělení datagramů dojde až podle IPv6 adresy na zařízení pracujícím na třetí, tedy síťové, vrstvě ISO/OSI modelu.

5.1.2 Multicast Listener Discovery (MLD)

Alternativou k IGMP je v prostředí IPv6 Multicast Listener Discovery (MLD) definovaný v RFC 2710. I přes odlišný název má ale stejnou funkci, tedy informování routerů o zájmu účastníků připojit se k odběru dat z daných multicastových skupin. MLDv1 přitom neuchovává informace o odesílatelích (není *source-specific*).

zpráva	popis
Multicast Listener Query (typ zprávy 130)	Ekvivalent k <i>IGMP Membership Query</i> zasílaného routerem ke odhalení posluchačů multicastových skupin na místní síti (obecný i specifický dotaz). Zprávu odesílá multicastový router na adresu všech MLDv2 uzlů na místní lince, tedy ff02::16, nebo adresu všech uzlů na lince, tedy ff02::1.
Multicast Listener Report (typ zprávy 131)	Ekvivalent k <i>IGMP Membership Report</i> , kterým posluchač signalizuje routeru svůj zájem o připojení se k dané multicastové skupině. Zprávu posluchač odesílá na adresu skupiny, jejíž data by rád odebíral.
Multicast Listener Done (typ zprávy 132)	Ekvivalent k <i>IGMP Leave Group</i> . Zpráva zasílaná routeru v případě opuštění skupiny posledním posluchačem na segmentu Zprávu odesílá posluchač na adresu ff02::2, tedy všechny směrovače na místní lince.

tab. 5 - zprávy Multicast Listener Discovery

Na zprávu 132 přitom směrovače mohou zareagovat dvěma způsoby [28]. V případě, že router obdrží *Multicast Listener Done* od jiného uzlu, než od kterého naposledy obdržel zprávu *Multicast Listener Report*, znamená to, že na dané skupinové adrese poslouchá alespoň jeden další uzel – router si tedy uchová aktivní *multicast forwarding state* pro danou skupinu a síťové rozhraní.

V případě, že *Multicast Listener Done* obdrží router právě od uzlu, který naposledy ohlašoval připojení ke skupině, může to znamenat, že na daném rozhraní už neexistuje žádný příjemce dané multicastové skupiny. Router proto iniciuje query zprávu typu 130 s cílem zjistit skutečný počet posluchačů dané skupiny. Okamžitá odpověď všech uzlů by ale mohla zahltit směrovač, proto si každý uzel nastaví náhodné zpoždění pro svou odpověď. Vzhledem k tomu, že odpověď odchází na adresu skupiny (typ 131), zachytí ji i ostatní uzly na místní síti. A protože by bylo další zasílání *Multicast Listener Reportů* zbytečné (routeru stačí jediný), svoji zpožděnou odpověď zahodí úplně.

Ekvivalentem k IGMPv3 [29] je v prostředí IPv6 MLDv2 (RFC 3810), které kvůli rozdílnému formátu hlavičky není kompatibilní s MLDv1. Nově přináší možnost filtrování vysílacích zdrojů (podpora *Source Specific Multicast*) prostřednictvím pravidel INCLUDE a EXCLUDE.

Zatímco INCLUDE signalizuje zájem posluchače o příjem paketů pouze od zdrojů definovaných ve zprávě, EXCLUDE signalizuje zájem posluchače o vyloučení všech zdrojů ve zprávě definovaných. Jednotlivá pravidla lze volně kombinovat a vytvářet tak základní logické operace jako průnik nebo sjednocení.

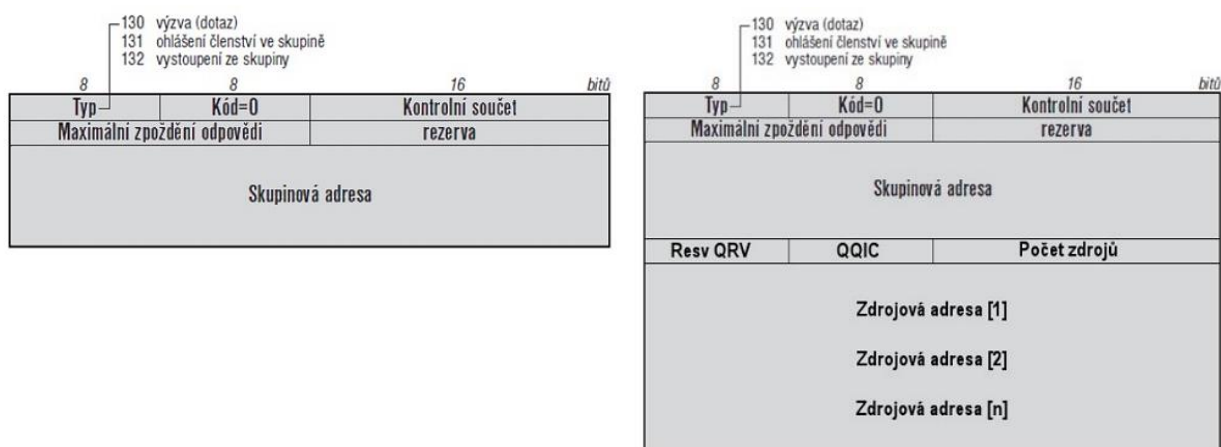
Queries mohou být celkem trojího druhu

- **General** – query dotaz routeru na posluchače skupin
- **Multicast address specific** – query pro specifickou skupinu
- **Multicast address and source specific** - query pro specifickou skupinu a zdroj

Stejně tak i *Reports*

- **Current state** (vyžádaný) – odpověď na query routeru, specifikuje include/exclude mód pro každou skupinovou adresu
- **Filter mode change** (nevyžádaný) – oznámení změny include/exclude módu pro jednu či skupinových adres
- **Source list change** (nevyžádaný) – oznámení o změně zdroje

Zpráva je zasílána všem MLDv2 směrovačům na síti (adresa ff02::16) a obsahuje více záznamů pro všechny využívané skupiny. Každý příjemce tedy odesílá jednu zprávu všem směrovačům, popisující jeho zájem odebrat různé skupiny.



obr. 70 – formát MLDv1 a MLDv2 [30]

5.2 Směrování multicastů v IPv6

Důležitou změnou v prostředí IPv6 je opuštění směrovacího protokolu PIM v módu Dense. Jeho využití v IPv4 je i dnes spíše jen teoretické, u IPv6 ale už chybí zcela. Dalším omezením je absence Cisco proprietárního *Auto-RP* módu. Využití protokolu PIM v módu Sparse ale vyžaduje, aby všechny zúčastněné routery znaly adresu RP. Toho je v prostředí IPv6 možno docílit několika cestami [31].

5.2.1 Statická konfigurace RP

Statická konfigurace IP adresy RP do všech routerů na síti je nepohodlná už v IPv4. Nový IP protokol toho nepohodlí svými dlouhými hexadecimálními adresami dále prohlubuje. Dalším problémem je jen malá odolnost vůči poruchám v případě statické konfigurace RP. Proto také vzniklo několik automatizovaných alternativ.

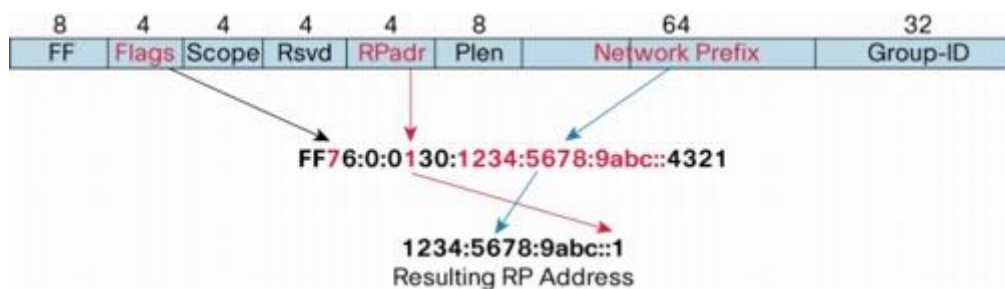
5.2.2 Mapovací agent BSR

Mapovací agent *BSR* je otevřenou alternativou k Cisco proprietárnímu *Auto-RP*, který si ale na rozdíl od *BSR* cestu do specifikace IPv6 našel. To na rozdíl od *Auto-RP* nevyužívá *PIM dense-mode* k zaplavení všech svých sousedů informacemi o kandidátech na *RP*. Místo toho využívá standardní PIM zprávy, které se šíří sítí od routeru k routeru (*hop-by-hop basis*). Pro správnou funkci *BSR* je nutné zvolit alespoň jednoho kandidáta na *BSR* router (označení *C-BSR*, vybírá *RP* kandidáty) a alespoň jednoho *RP* kandidáta (označení *C-RP*), ideálně ale samozřejmě více kandidátů tak, aby *BSR* router mohl v případě nedostupnosti prvního *RP* kandidáta zvolit jeho náhradníka. Obyčejně jsou *BSR* routery sami zároveň i *RP* kandidáty.

C-RP vysílá periodicky tzv. *candidate-RP-advertisement* (*C-RP-Adv*) zprávy adresované na unicastovou adresu zvoleného *BSR* routeru. V nich oznamují *BSR* routeru jejich zájem stát se *RP*. *C-RP-Adv* zpráva obsahuje adresu propagovaného *C-RP* a volitelně i seznam skupinových adres s informací o délce masky (pro určení prefixu sítě). *BSR* router pak vytvoří seznam těchto *C-RPs* a periodicky je odesílá všem směrovačům na síti jako tzv. *bootstrap messages* (*BSMs*).

5.2.3 Embedded RP

Nejelegantnější variantou je pak bezesporu tzv. *Embedded RP* [32]. Široký adresní rozsah IPv6 umožňuje do IP adresy multicastové skupiny vložit také IP adresu *Rendezvous Point* (RP), jehož ustanovení je podmínkou pro funkci PIM-SM. Skupinová adresa obsahující adresu RP má nastaveny příznaky R, P a T z části *Flags* na 1 a začíná prefixem **ff7::/12**.



obr. 71 - vtisknutí *Rendezvous Point* do IPv6 adresy (RFC 3956) [33]

128bit IPv6 adresa je v případě použití "vtisknutí" adresy RP rozdělena následovně:

- 8bit pole FF u *Embedded RP* vždy prefix "1111"
- 4bit pole Flags u *Embedded RP* vždy "0111"
- 4bit pole Scope standardní určení dosahu paketu (lokální, globální, ...)
- 4bit pole Rsvd u *Embedded RP* vždy "0000"
- 4bit pole RPadr určuje adresu rozhraní RP
- 8bit pole Plen určuje nastavení délky prefixu sítě z následující položky
- 64bit pole Network p. prefix sítě
- 32bit pole Group ID ID multicastové skupiny

Další možností je využití varianty PIM-SM, tzv. Bidirectional PIM (BIDIR-PIM, popsán v minulé kapitole), nebo systému *Source Specific Multicast* (SSM). Podobně jako u IPv4, tak i v případě IPv6 SSM pro svou funkci nevyžaduje definování adresy RP – klient vytváří vlastní *Shortest Path Tree* distribuční stromy přímo k multicastovému zdroji.

5.3 Simulace směrování IPv6 multicastu v prostředí Cisco IOS

Pro demonstraci reálné funkce směrování multicastových datagramů v prostředí IPv6 bylo vybráno prostředí Cisco IOS [34], jelikož se jedná de facto o dnešní průmyslový standard na poli telekomunikací.

5.3.1 Platforma

Pro analýzu a demonstraci funkce síťové infrastruktury bývá v prostředí Cisco IOS často využíván program *Packet Tracer*. Jeho možnosti se ale pro simulaci směrování multicastů v prostředí IPv6 ukázaly jako příliš omezené. *Packet Tracer* totiž umožňuje simulovat pouze omezenou sadu příkazů IOS, která v sobě nezahrnuje podporu směrování multicastů ani v IPv4, natož v IPv6.

Jako alternativa byl zvolen simulátor *Graphic Network Simulator 3* (GNS3), který je k dispozici zdarma pod licencí GPLv2. GNS3 je souborem několika různých programů. Nejdůležitější z nich je pravděpodobně *Dynamips*, který obstarává emulaci routerů řady 1700, 2600, 3600, 3700 a 7200 společnosti Cisco. *Dynamips* na rozdíl od *Cisco Packet Traceru* není simulátor, ale skutečně emulátor původních Cisco obrazů IOS, které zavádí do paměti vytvořených virtuálních strojů. Na použití těchto obrazů je zapotřebí oficiální licence firmy Cisco. Uživatel tak může operovat s kompletní sadou příkazů použité verze IOS.

5.3.2 Verze IOS a HW vybavení

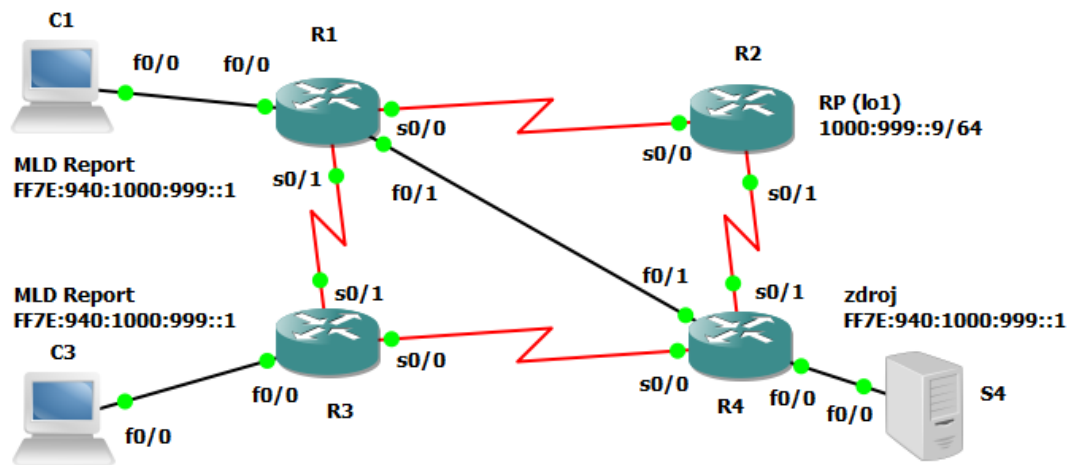
Funkce směrování IPv6 multicastů byla simulována na síťové infrastruktuře sestávající se celkem ze sedmi routerů Cisco 2691 s přídatnými moduly WIC-1T (propojení přes sériovou linku standardem V.35), z nich dva byly použity jako klienti odebírající multicastová data a jeden jako jejich zdroj. Vybavení je to dnes už spíše muzejní, ale pro provedení funkční simulace dostačující. Všechny zařízení pracovaly s IOS ve verzi 12.4(25d) ze srpna 2010.

```
Cisco IOS Software, 2600 Software (C2691-ADVENTERPRISEK9-M), Version 12.4(25d),
RELEASE SOFTWARE (fcl) Technical Support: http://www.cisco.com/techsupport
Copyright (c) 1986-2010 by Cisco Systems, Inc.
Compiled Wed 18-Aug-10 05:35 by prod_rel_team
```

```
Cisco 2691 (R7000) processor (revision 0.1) with 124928K/6144K bytes of memory.
Processor board ID XXXXXXXXXXXX
R7000 CPU at 160MHz, Implementation 39, Rev 2.1, 256KB L2, 512KB L3 Cache
2 FastEthernet interfaces
DRAM configuration is 64 bits wide with parity enabled.
55K bytes of NVRAM.
16384K bytes of ATA System CompactFlash (Read/Write)
```

obr. 72 - výpis *show version* použitého routeru Cisco 2691

5.3.3 Síťová topologie a nastavení prvků



obr. 73 - síťová topologie pro simulaci směrování IPv6 multicastů

zařízení	fo/0	fo/1	so/0	so/1	výchozí brána / lo
S4	1000:7::2/64	-	-	-	1000:7::1/64
C1	1000:6::2/64	-	-	-	1000:6::1/64
C2	1000:8::2/64	-	-	-	1000:8::1/64
R1	1000:6::1/64	1000:5::1/64	1000:1::2/64	1000:2::2/64	-
R2 (RP)	-	-	1000:1::1/64	1000:3::1/64	1000:999::9/64 (lo)
R3	1000:8::1/64	-	1000:4::1/64	1000:2::1/64	-
R4	1000:7::1/64	1000:5::2/64	1000:4::2/64	1000:3::2/64	-

tab. 6 - nastavení jednotlivých síťových prvků a jejich rozhraní

- jednotlivé podsítě byly ostatním směrovačům propagovány přes OSPFv3
- na instalovaných směrovačích byl spuštěn multicastový směrovací démon PIM-SM
- klienti C1 a C2 oba vysílali MLD Report na adresu zvolené skupiny (ID 1) s vloženou IPv6 adresou Rendezvous point (loopback rozhraní routeru R2)
- multicastový provoz byl generován serverem S4 jako posloupnost ICMPv3 zpráv adresovaných na IPv6 adresu zvolené multicastové skupiny

Podrobná konfigurace všech prvků viz. příloha č. 2.

5.3.4 Určení a nastavení Embedded RP

Obecný popis vtisknutí Rendezvous Point IPv6 adresy do skupinové IPv6 adresy [35] už byl detailně popsán v jedné z minulých kapitol. Konkrétně pro případ simulace by byl postup následující.

1. IPv6 adresa loopback rozhraní Rendezvous Point - **1000:999::9/64**
2. z této adresy:
 - a. RP rozhraní ID **9 (::9)**
 - b. délka prefixu /64, tedy **40** hexadecimálně
 - c. RP prefix **1000:999**
3. určení dosahu adresy (*scope*)
 - a. nabývá hodnot **1, 2, 4, 5, 8** a **E**, kde **E** značí *Global*
4. kompletace výsledné adresy podle vzoru

FF7<Scope>:0<RP rozhraní ID><délka prefixu v HEX>:<64bit RP prefix>:<32bit ID rozhraní>

5. výsledný rozsah skupinových IPv6 adres s vtisknutou adresou RP je tedy **FF7E:940:1000:999::1 až FF7E:940:1000:999::FFFF:FFFF**

Na směrovačích tedy není nutné provést jakoukoliv statickou konfiguraci IPv6 adresy RP, jelikož je automaticky odvozována z cílových IPv6 adres probíhajících paketů. Posledním krokem je definování adresy RP na samotném RP a vytvoření *access-listu*, který umožní RP přijímat multicastové skupiny s různým ID rozhraní odvozené od adresy RP (tedy ve formátu FF7E:940:1000:999::X).

Simulace připojení uzlu k multicastové skupině je na routeru provedena jako MLD report. Unicastové směrování musí být vypnuto a veškerý provoz nasměrován na výchozí bránu.

<pre>! ipv6 unicast-routing ipv6 multicast-routing ! ipv6 pim rp-address 1000:999::9 RPLIST ipv6 access-list RPLIST permit ipv6 any ff7e:940:1000:999::/96</pre>	<pre>! no ip routing ! interface FastEthernet0/0 ipv6 address 1000:6::2/64 ipv6 mld join-group FF7E:940:1000:999::1 ! ipv6 route ::/0 1000:6::1</pre>
--	---

obr. 74 - specifické nastavení RP a C1 (C2)

5.3.5 Analýza provozu

Prostředí systému IOS společnosti Cisco nabízí velmi rozsáhlé možnosti analýzy IPv6 provozu.

```
R4#debug ipv6 ?
  access-list      IPv6 access list debugging
  cef              IPv6 CEF information
  cpc             IPv6 Common Parsing Cache debugging
  dhcp           IPv6 DHCP debugging
  icmp          ICMPv6 debugging
  inspect       Stateful inspection events
  interface     IPv6 interface debugging
  mfib         IP Multicast forwarding information base
  mld         Multicast Listener Discovery
  mobile       MIPv6 Debugging
  mrib         Multicast Route DB
  nat          NAT-PT events
  nd           IPv6 Neighbor Discovery debugging
  ospf        OSPF information
  packet      IPv6 packet debugging
  pim       Protocol Independent Multicast
  policy      IPv6 policy-based routing debugging
  pool        IPv6 prefix pool debugging
  port-mapping IPv6 PAM events
  rip         RIP Routing Protocol debugging
  routing     IPv6 routing table debugging
  virtual-reassembly IPv6 Virtual Fragment Reassembly (VFR) debugging
```

obr. 75 - možnosti analýzy IPv6 systému IOS

Klíčový je zejména *debug ipv6 pim* pro analýzu funkce protokolu PIM, *debug ipv6 mld* pro analýzu funkce MLD a samozřejmě také *show ipv6 mroute* pro zobrazení stavu distribučních stromů na daném routeru.

Před spuštěním multicastového zdroje není na *designated routeru* zdroje vygenerován žádný distribuční strom.

```
R4#show ipv6 mroute
No mroute entries found.
```

Při generování takového provozu opakovaným zasíláním ICMPv3 zprávy na zvolenou multicastovou adresu *FF7E:940:1000:999::1* se na DR multicastového zdroje (R4) vytvoří první Shortest Path Tree distribuční strom. Zdrojem je *1000:7::2* (tedy sám multicastový zdroj) a skupinovou adresou právě *FF7E:940:1000:999::1*. Na dotazy přitom odpovídají oba posluchači multicastové skupiny. Vytvořený strom má příznaky SFT, tedy S – Sparse mode, F – zasílání registračních zpráv a T – jedná se o Shortest Path Tree.

```

S4#ping FF7E:940:1000:999::1 repeat 3
Output Interface: fastethernet0/0
Type escape sequence to abort.
Sending 3, 100-byte ICMP Echos to FF7E:940:1000:999::1, timeout is 2 seconds:
Packet sent with a source address of 1000:7::2

Reply to request 0 received from 1000:6::2, 172 ms
Reply to request 0 received from 1000:8::2, 224 ms
Reply to request 1 received from 1000:8::2, 104 ms
Reply to request 1 received from 1000:6::2, 176 ms
Reply to request 2 received from 1000:8::2, 20 ms
Reply to request 2 received from 1000:6::2, 68 ms
...
Success rate is 100 percent (3/3), round-trip min/avg/max = 20/112/224 ms
6 multicast replies and 0 errors.

```

obr. 77 - generování multicastového provozu prostřednictvím ICMPv3 zpráv

```

R4#show ipv6 mroute
Multicast Routing Table

...

(1000:7::2, FF7E:940:1000:999::1), 00:00:10/00:03:19, flags: SFT
  Incoming interface: FastEthernet0/0
  RPF nbr: 1000:7::2
  Immediate Outgoing interface list:
    FastEthernet0/1, Forward, 00:00:10/00:03:19
    Serial0/0, Forward, 00:00:10/00:03:19

```

obr. 76 - vytvořený SPT distribuční strom na DR multicastového zdroje

Snižující se latence ilustrují přepnutí ze sdíleného distribučního stromu s neoptimální cestou přes RP (Rendezvous Point Tree) na nejkratší možný distribuční strom (Shortest Path Tree).

```

R4#debug ipv6 pim
IPv6 PIM debugging is on
R4#
IPv6 PIM: [XXXXXXXXXX] (1000:7::2, FF7E:940:1000:999::1/128) MRIB update (t=1)
IPv6 PIM: [XXXXXXXXXX] (1000:7::2, FF7E:940:1000:999::1/128) FastEthernet0/0 MRIB
update (f=2A, c=20)
IPv6 PIM: [XXXXXXXXXX] (1000:7::2, FF7E:940:1000:999::1) Signal present on
FastEthernet0/0
IPv6 PIM: [XXXXXXXXXX] (1000:7::2, FF7E:940:1000:999::1/128) FastEthernet0/0 MRIB
modify !NS
IPv6 PIM: (1000:7::2, FF7E:940:1000:999::1) Start registering to 1000:999::9
IPv6 PIM: (1000:7::2, FF7E:940:1000:999::1) Tunnel0 J/P state changed from Null to
Join
IPv6 PIM: (1000:7::2, FF7E:940:1000:999::1) Tunnel0 FWD state change from Prune to
Forward
IPv6 PIM: [XXXXXXXXXX] (1000:7::2, FF7E:940:1000:999::1/128) Tunnel0 MRIB modify F

```

obr. 78 - router informuje o novém zdroji a směruje provoz na RP

```

R1#show ipv6 mroute
Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group,
       C - Connected, L - Local, I - Received Source Specific Host Report,
       P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
       J - Join SPT
Timers: Uptime/Expires
Interface state: Interface, State

(*, FF7E:940:1000:999::1), 00:05:00/00:02:30, RP 1000:999::9, flags: SCJ
  Incoming interface: Serial0/0
  RPF nbr: FE80::C205:11FF:FE50:0
  Immediate Outgoing interface list:
    Serial0/1, Forward, 00:04:55/00:02:30
    FastEthernet0/0, Forward, 00:05:00/never

(1000:7::2, FF7E:940:1000:999::1), 00:01:16/00:02:12, RP 1000:999::9, flags: SRJ
  Incoming interface: Serial0/0
  RPF nbr: FE80::C205:11FF:FE50:0
  Immediate Outgoing interface list:
    Serial0/1, Null, 00:01:16/00:02:30
  Inherited Outgoing interface list:
    FastEthernet0/0, Forward, 00:05:00/never

(1000:7::2, FF7E:940:1000:999::1), 00:01:18/00:02:12, flags: SJT
  Incoming interface: FastEthernet0/1
  RPF nbr: FE80::C203:11FF:FE50:1
  Inherited Outgoing interface list:
    Serial0/1, Forward, 00:04:55/00:02:30
    FastEthernet0/0, Forward, 00:05:00/never

```

obr. 8o - vytvořené distribuční stromy na R1 (DR pro C1)

Výpis vytvořených distribučních stromů na routeru R1, který je zároveň designated routerem pro C1 ukazuje vytvořený sdílený strom, který se rozšiřuje z rozhraní vedoucího k RP (serialo/o) jednak směrem k C1 (fao/o), jednak směrem k C2 přes R3 (so/1). Příznak C sdíleného stromu označuje přímé připojení posluchače multicastové skupiny. Příznak T distribučního stromu specifického zdroje ukazuje, že ke směrování datagramů je využíváno právě tohoto stromu.

```

R1#show ipv6 mld groups
MLD Connected Group Membership
Group Address                               Interface      Uptime      Expires
FF7E:940:1000:999::1                       FastEthernet0/0 00:46:34   00:03:32

```

obr. 79 - posluchači v MLD tabulce routeru R1

5.4 Možnosti implementace v síti lokálního ISP

5.4.1 Stav alokace IPv4 a cesty řešení

Stav alokace IPv4 prostoru a rychlost jeho čerpání už dávno dával tušit, že reálně není možné, aby měl každý uživatel internetu přidělenou vlastní veřejnou IPv4 adresu. Přestože 32bit IPv4 rozsah nabízí možnost teoreticky adresovat až 2^{32} (tedy 4 294 967 296) adres, mnohé z nich jsou rezervovány pro experimentální a jiné využití (konkrétně 588 514 304 adres podle RFC 6052). Další, často i celé /8 bloky (celý IPv4 prostor se dělí na 256 těchto bloků) jsou historicky přiděleny společnostem jako Apple nebo Microsoft a nemohou tedy být použity pro veřejné adresování. Počet internetových uživatelů se přitom v roce 2011 odhadoval asi na 2,5 miliardy celosvětově. Není proto divu, že organizace IANA (Internet Assigned Numbers Authority), která celosvětově spravuje alokaci IP rozsahů, již 1. ledna 2011 oznámila, že vyčerpala rozsah alokovatelných IPv4 adres přidělovaných místním správčům jako je např. evropská organizace RIPE NCC.

Možnosti, jak získat veřejný IPv4 rozsah, existují v zásadě dvě:

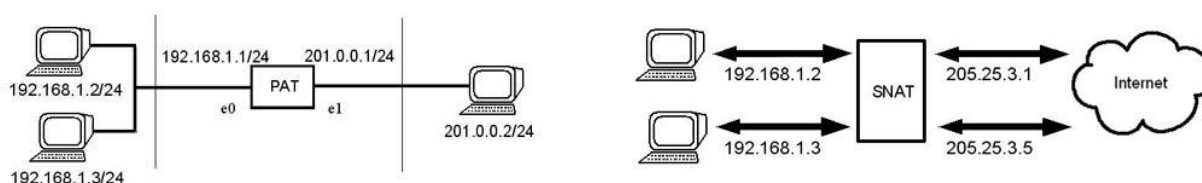
1. Stát se tzv. LIR, tedy členem sdružení lokálního registrátora, jímž je v Evropě již zmíněný RIPE NCC. Členové mají v současné době nárok na přidělení /22 IPv4 rozsahu (tedy 1 024 adres) za podmínky, že již mají přidělený i IPv6 rozsah (který RIPE přiděluje všem členům automaticky). Členství v RIPE NCC je ovšem zpoplatněno částkou 1800 EUR ročně, proto členství není vždy samozřejmostí, především pro menší poskytovatele. Výhodou je ale nezávislost přiděleného IP rozsahu a možnost s ním libovolně nakládat.
2. Získat veřejný IPv4 rozsah od svého upstream providera (který je již sám členem RIPE NCC). V případě, že internetový poskytovatel získává IP konektivitu od upstream providera, musí k ní nutně získat i veřejný IP rozsah. Konkrétní nabídky se mohou lišit, běžně to ale bývá /24 (tedy 254 alokovatelných adres). Tyto adresy zůstávají přiděleny upstream providerovi, při změně dodavatele konektivity tedy nutně dochází k přečíslování veřejných adres poskytovatele, což přináší řadu dalších problémů.

V obou případech je ale poskytovatel nucen vystačit s omezeným IP rozsahem, který má k dispozici. Drobnou pomoc představuje koncept dynamické alokace IP adresy, kdy veřejná IP adresa není vázána na konkrétní zařízení klienta, ale je přidělována dynamicky principem *first-*

come, first-served z poolu veřejných adres. Tento postup se ale neukázal být příliš populárním, jelikož střídání přidělené IP adresy přinášelo klientům řadu komplikací.

Jediným efektivním řešením tak zůstává využití překladu adres, nejčastěji na výchozí bráně poskytovatele, tedy techniky *Network Address Translation* (NAT). Jakkoliv je to mnohými démonizovaný koncept, který narušuje TCP/IP vlastní tzv. *end-to-end principle* (tedy vyvarování se implementace technik, které na dané vrstvě nijak nezlepšují její parametry), se snižujícím se počtem alokovatelných IPv4 adres dochází a bude docházet k jeho využití stále častěji.

Díky překladu adres tak může internetový poskytovatel adresovat své klienty na jednom z privátních rozsahů, z nichž ten největší, 10.0.0.0/8 čítá celkem 16 777 214 adresovatelných síťových uzlů. Tisíce privátních IP adres je tak překládáno třeba i na jedinou veřejnou IP adresu s různými porty, jedná se o tzv. *Port Address Translation* (PAT). Internetový poskytovatel přitom může přistoupit i k překladu 1:1, tedy využít tzv. SNAT, který se vyznačuje pevným spojením veřejné s neveřejnou IP.



obr. 81 – PAT a SNAT [36]

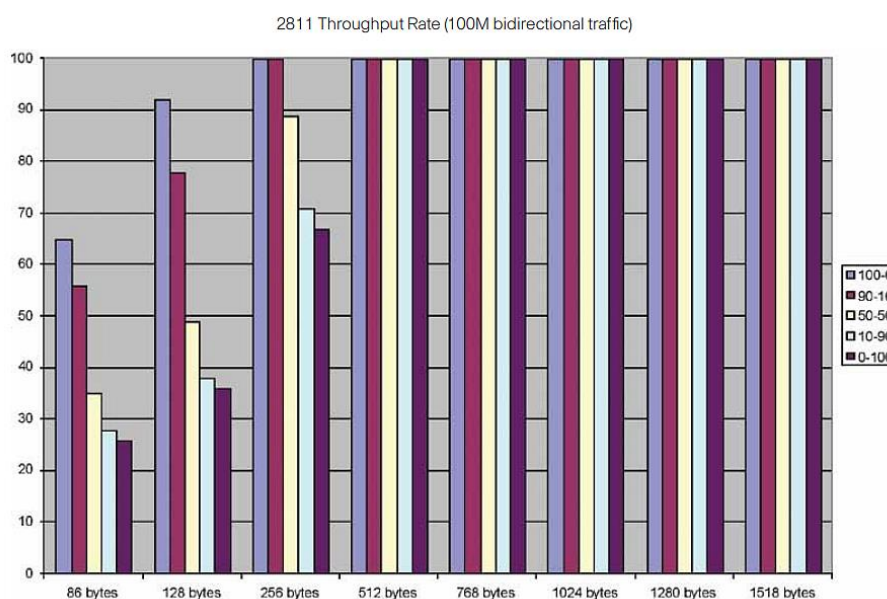
5.4.2 Kompatibilita IPv6 multicastu se síťovými prvky

I po vyčerpání alokovatelného IPv4 rozsahu tedy internetoví poskytovatelé nebudou nuceni hromadně přecházet na novou verzi internetového protokolu. Dalším problémem bývá často kompatibilita nasazených zařízení.

Byť není myšlenka IPv6 zrovna mladá (nový IP protokol byl definován poprvé už v roce 1995 v RFC 1883), jeho podpora ve spojení s vícesměrovým vysíláním není v HW dodnes samozřejmostí. Např. společnost Cisco uvedla podporu IPv6 multicastu v IOS verze 12.3, která byla uvedena v roce 2005. V případě Linuxu byla podpora nového IP protokolu uvedena již mnohem dříve v kernelu 2.2 z roku 1998. Všechny moderní distribuce tedy IPv6 podporují. K směrování IPv6 multicastu v prostředí Linuxu je zapotřebí instalace externího směrovacího démonu, nejčastěji *mrd6*, jehož vývoj začal v roce 2004. Kompatibilita síťových prvků operujících na druhé vrstvě ISO/OSI modelu je diskutována podrobně v jedné z dalších kapitol.

Operace se čtyřnásobně dlouhými adresami (32 bit vs. 128 bit) navíc nadměrně zatěžují HW, při vývoji síťového prvku je s tímto faktem nutné počítat. Při nedostatečně výkonnostně dimenzovaném zařízení pak podpora nového IP protokolu nemůže být dodatečně přidána pouhou aktualizací SW vybavení prvku.

Graf níže [37] zobrazuje propustnost routeru Cisco 2811 při souběžném směrování IPv4 a IPv6 provozu procházejícího přes obě jeho 100Mbit rozhraní. Zatímco fialový sloupec značí všechny pakety adresované jako IPv6, modrý graf pak všechny pakety adresované jako IPv4 a žlutý sloupec pakety rozložené v poměru 50:50.



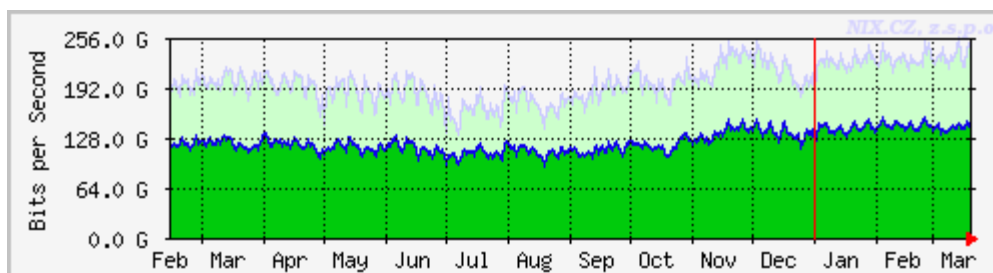
obr. 82 – srovnání náročnosti směrování IPv4 a IPv6

Z grafů je patrné, že při směrování malých paketů (velikost 86B) dochází k výraznému poklesu výkonnosti routeru při adresování IPv6. Propustnost se snižuje z 65Mbit/s až na 27Mbit/s. Podobným poměrem se i prodlužují latence a roste zátěž procesoru. Spolu s délkou IP adresy totiž roste i čas potřebný na její zpracování. Záznamy v směrovacích tabulkách se prodlužují a vyhledávání v ní proto zpomaluje.

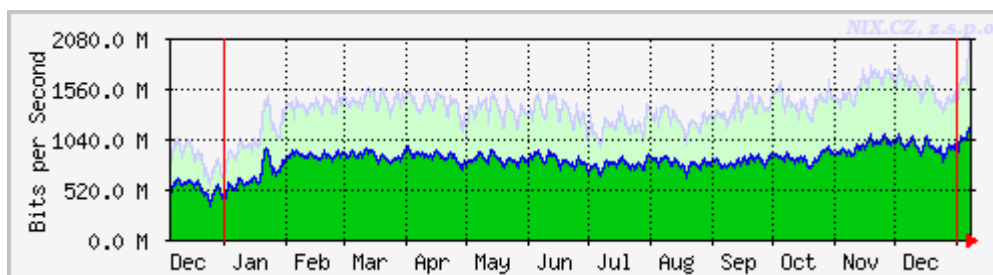
Hlavní problém ale představuje rostoucí délka záznamů v směrovací tabulce. Tyto jsou uloženy v paměti, která je běžně nazývána jako TCAM a obvykle bývá omezena na 512k záznamů v případě použití IPv4 adresace. V případě IPv6 adresace se tato kapacita snižuje na 256k záznamů (byť poměr neodpovídá). Kvůli velkému rozsahu a absenci překladu adres se předpokládá výrazná fragmentace propagovaných sítí, a tedy i rychlý růst počtu záznamů v globální směrovací IPv6 tabulce.

5.4.3 Aplikace a budoucnost

Rozšíření samotného vícesměrového vysílání v prostředí IPv4 není velké a omezuje se z velké části na izolované místní aplikace. Přenosy adresované jako IPv6 přitom v dnešní době představují asi 1% z celkového internetového provozu. Spojení obou přístupů je proto logicky ojedinělé a v praxi na něj snad ani nelze narazit, kromě testů probíhajících na akademické půdě, apod.



obr. 84 – IPv4 přenosy NIX.CZ (do 21.3.2013)



obr. 84 – IPv6 přenosy NIX.CZ (do 21.3.2013)

Ba co více, vzhledem k tomu, že byla pro multicast vyhrazena celá adresní třída D, nehrozí jí ani bezprostřední vyčerpání jako v případě unicastového vysílání v prostředí IPv4. Jinými slovy řečeno, multicastové vysílání je na stavu alokace unicastových IPv4 adres nezávislé. Proto také nejsou provozovatelé nuceni k přechodu na IPv6, alespoň ne kvůli vícesměrovému vysílání.

ZÁVĚR

Přestože koncept vícesměrového vysílání přináší zvýšenou efektivitu přenosu dat, dosud se nedočkal širšího nasazení. Jako každá nová technologie, která není zpětně kompatibilní se svým předchůdcem, musel i multicast vyčkat na vhodnou aplikaci, která by si jeho využití vyžádala. Až do nedávné doby byly totiž vzrůstající požadavky na přenosovou kapacitu počítačových sítí dostatečně pokryty zvyšováním jejich propustnosti na fyzické vrstvě. Zvrat ale přišel s nástupem služeb jako Internet Protocol TV a Video on Demand, kde lineární škálování dosavadního principu jednosměrového vysílání nedokázalo uspokojit rychle rostoucí požadavky na přenosovou kapacitu. Právě vícesměrové vysílání se ukázalo být východiskem, obecným principem, který replikací datagramů umožnil dále zvýšit přenosovou kapacitu na síťové vrstvě při zachování přenosové rychlosti na vrstvě fyzické.

Cílem této diplomové práce bylo nejen teoreticky popsat princip vícesměrového vysílání, ale také demonstrovat jeho výhody na vhodné aplikaci, kterou se stala realizace služby IPTV v síti lokálního internetového poskytovatele.

V teoretické části byly popsány všechny specifika vícesměrového vysílání, včetně principu adresování multicastových datagramů a konstrukce specifických i sdílených distribučních stromů. Důraz byl kladen na rozbor fungování nejrozšířenějšího multicastového směrovacího protokolu Protocol Independent Multicast v režimu Sparse ve všech jeho stádiích, stejně tak i protokolu Internet Group Management Protocol, který rozšiřuje implementaci IPv4 o podporu IP multicastu. Pro úplnost byly dále zmíněny i ostatní běžně používané způsoby vysílání a adresování IP paketů.

Praktická část práce pojednává zejména o využití vícesměrového vysílání pro přenosy IPTV, definuje základní kvalitativní požadavky na přenos tohoto obsahu a popisuje konfiguraci všech potřebných síťových prvků včetně nastavení politik Quality of Service. Důležitou součástí je také detailní analýza a popis chování multicastového provozu v jeho různých stádiích prováděná nástrojem tcpdump na několika sondách v různých částech síťové topologie.

Další část srovnává výkonnostní škálování vícesměrového s jednosměrovým vysíláním a popisuje důsledky tohoto škálování na reálné síťové aplikace. Poslední část diplomové práce shrnuje fungování vícesměrového vysílání v prostředí IPv6, včetně provedení praktické simulace v prostředí Graphical Network Simulator na síti sestávající se z multicastového zdroje, několika Cisco směrovačů a klientů odebírajících data ze skupinových adres.

CONCLUSION

Although the concept of multicast delivers increased efficiency of data transfer, it has not achieved widespread. Like any new technology that is not backward compatible with its predecessor, multicast had had to wait for the appropriate application that would have justify its implementation. Until recently, the increasing demand for bandwidth in computer networks was adequately covered by increasing network's throughput at the physical layer. Turning point came with the advent of services such as Internet Protocol TV and Video on Demand, where the linear scaling of existing unicast principle fails to meet the rapidly growing demand for bandwidth. Multicast has proven to be the starting point, the general principle that by the datagram replication allows to further increase the bandwidth of the network layer while keeping the bit rate at the physical layer.

The aim of this thesis was not only theoretically describe the principle of multicast, but also demonstrate its advantages on the appropriate application, which has become the realization of IPTV services in a network of local Internet provider.

The theoretical part describes all the specifics of multicast, including multicast datagrams addressing and design of Specific and Shared Distribution Trees. Emphasis was placed on the analysis of the most popular multicast routing protocol Protocol Independent Multicast operating in Sparse Mode in all its stages, as well as the Protocol Internet Group Management Protocol, which extends the IPv4 implementation to support IP multicast. For completeness, other commonly used methods of transmission and addressing IP packets have been mentioned.

The practical part of the thesis is focused on the use of multicast in IPTV application, defines the basic quality requirements for the transmission of this content and describes the configuration of all network elements and implementing Quality of Service policies. The important part is a detailed analysis and description of the behavior of multicast traffic in its various stages by tcpdump tool running on several probes placed in different parts of the network topology.

Another part compares the performance scaling of unicast and multicast transmission and describes the implications of this scaling on real network applications. The last part of the thesis summarizes the function of multicast IPv6 environment, including the implementation of practical simulation in Graphical Network Simulator on a network consisting of a multicast source, several Cisco routers and multicast address listeners.

SEZNAM POUŽITÉ LITERATURY

- [1] Mapping IP Multicast to MAC-Layer Multicast. *technet.microsoft.com* [online]. ? [cit. 2013-05-11]. Dostupné z: <http://technet.microsoft.com/en-us/library/cc957928.aspx>
- [2] WILLIAMSON, Beau. Developing IP multicast networks. 5. print. Indianapolis, Ind: Cisco, 2000. ISBN 15-787-0077-9.
- [3] IANA Guidelines for IPv4 Multicast Address Assignments. In: M. COTTON a L. VEGODA. *Internet Engineering Task Force (IETF)* [online]. 2010 [cit. 2013-05-14]. Dostupné z: <http://tools.ietf.org/html/rfc5771>
- [4] EDWARDS, Brian M, Leonard A GIULIANO a Brian R WRIGHT. Interdomain multicast routing: practical Juniper Networks and Cisco systems solutions. Boston: Addison-Wesley, c2002, xxiii, 356 p. ISBN 02-017-4612-3.
- [5] B. FENNER, M. HANDLEY, H. HOLBROOK a I. KOUVELAS. RFC 4601: Protocol Independent Multicast – Sparse Mode (PIM-SM): Protocol Specification (Revised). In: Ietf.org: Network Working Group [online]. AT&T / Cisco Systems / UCL, 2006 [cit. 2013-03-26]. Dostupné z: <http://www.ietf.org/rfc/rfc4601.txt>
- [6] ADAMS, J. NICHOLAS a W. SIADAK. RFC 3973: Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised). RFC 3973 [online]. 2005, č. 2005, s. 61 [cit. 2013-05-08]. Dostupné z: <http://tools.ietf.org/html/rfc3973>
- [7] PIM Sparse-Mode / Bootstrap Router (BSR) part II. BANK, Yuri. *Networking Adventures: Network protocols, design, and implementation* [online]. 2011. vyd. 2011 [cit. 2013-05-09]. Dostupné z: <http://yuri.easytospell.net/?p=426>
- [8] M. HANDLEY, I. KOUVELAS a T. SPEAKMAN. Bidirectional Protocol Independent Multicast (BIDIR-PIM). In: *Internet Engineering Task Force (IETF)* [online]. UCL, Cisco, 2007 [cit. 2013-05-14]. Dostupné z: <http://tools.ietf.org/html/rfc5015>
- [9] H. HOLBROOK a B. CAIN. Source-Specific Multicast for IP. In: *Internet Engineering Task Force (IETF)* [online]. Arastra, Inc., 2006 [cit. 2013-05-14]. Dostupné z: <http://tools.ietf.org/html/rfc4607>
- [10] Securing IP Multicast Services in Triple-Play and Mobile Networks. CISCO. Cisco.com [online]. ? [cit. 2013-05-09]. Dostupné z: <http://goo.gl/WbcXR>

- [11] SAVOLA, P. Overview of the Internet Multicast Routing Architecture. Ietf.org: Network Working Group [online]. T Category: Informational Jan, 2008 [cit. 2013-03-03]. Dostupné z: <http://tools.ietf.org/pdf/rfc5110.pdf>
- [12] [29] H. HOLBROOK, ARASTRA, INC., B. CAIN a B. HABERMAN. RFC 4604: Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source-Specific Multicast. In: *Ietf.org: Network Working Group* [online]. The Johns Hopkins University Applied Physics Laboratory, 2006 [cit. 2013-01-20]. Dostupné z: <http://tools.ietf.org/html/rfc4604>
- [13] Internet Group Management Protocol (IGMP). CISCO. *Technet.microsoft.com* [online]. 2005 [cit. 2013-05-09]. Dostupné z: [http://technet.microsoft.com/en-us/library/cc787925\(v=ws.10\).aspx](http://technet.microsoft.com/en-us/library/cc787925(v=ws.10).aspx)
- [14] [21] [22] QoS Introduction & Technology White Paper. *H3c.com* [online]. 2003-2013 [cit. 2013-05-09]. Dostupné z: <http://goo.gl/f8PMm>
- [15] IP Multicast Deployment Fundamentals. CISCO. *Cisco.com* [online]. ? [cit. 2013-05-09]. Dostupné z: http://www.cisco.com/en/US/tech/tk828/tech_brief09186a00800e9952.html
- [16] C. PARTRIDGE, W. MILLIKEN a T. MENDEZ. RFC 1546: Host Anycasting Service. In: Ietf.org: Network Working Group [online]. BBN, 1993 [cit. 2013-03-26]. Dostupné z: <http://www.ietf.org/rfc/rfc1546.txt>
- [17] NAVAS, Julio C. a Tomasz IMIELINSKI. GeoCast – Geographic Addressing and Routing. In: <http://www.comp.nus.edu.sg> [online]. Rutgers, The State University, 1997 [cit. 2013-03-03]. Dostupné z: <http://www.comp.nus.edu.sg/~bleong/geographic/related/navas97geocast.pdf>
- [18] Geocast. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-05-10]. Dostupné z: <http://en.wikipedia.org/wiki/Geocast>
- [19] [20] MINOLI, Daniel. IP multicast with applications to IPTV and mobile DVB-H. Hoboken: John Wiley, 2008, xvi, 357 s. ISBN 0470258152.
- [23] SDP: Session Description Protocol. In: Network Working Group, ietf.org [online]. University of Glasgow: ietf.org, 2006 [cit. 2013-05-08]. Dostupné z: <http://tools.ietf.org/html/rfc4566>
- [24] Internet Protocol version 6 (IPv6) Conformance and Performance Testing. *Net130.com* [online]. 2005 [cit. 2013-05-10]. Dostupné z: <http://goo.gl/NXOzz>

- [25] R. HINDEN a S. DEERING. RFC 4291: IP Version 6 Addressing Architecture. In: Ietf.org: Network Working Group [online]. Nokia / Cisco Systems, 2006 [cit. 2013-03-26]. Dostupné z: <http://www.ietf.org/rfc/rfc4291.txt>[27] Multicast Introduction. *H3c.com* [online]. ? [cit. 2013-05-10]. Dostupné z: <http://goo.gl/tRvq8>
- [28] Analýza protokolu MLD pro multicasting v IPv6 sítích. In: HANYÁŠ, Petr a Jiří KREJČÍŘ. Analýza protokolu MLD pro multicasting v IPv6 sítích [online]. 1. vyd. Vysoká škola báňská – Technická univerzita Ostrava, 2010 [cit. 2013-04-04]. Dostupné z: <http://goo.gl/RM4U1>
- [30] Analýza protokolu MLD pro multicasting v IPv6 sítích. HANYÁŠ, Petr a Jiří KREJČÍŘ. *Http://wh.cs.usb.cz* [online]. 2010 [cit. 2013-05-10]. Dostupné z: <http://goo.gl/RM4U1>
- [31] Implementing IPv6 Addressing and Basic Connectivity. In: Implementing IPv6 Addressing and Basic Connectivity [online]. 2001, 22.7.2011 [cit. 2013-04-04]. Dostupné z: <http://goo.gl/76Hwo>
- [32] LOCKER, Andreas a Hannes PAYER. Linux IPv6 Multicast Routing. In: Linux IPv6 Multicast Routing [online]. Department of Computer Science, University of Salzburg, 2005 [cit. 2013-04-04]. Dostupné z: <http://goo.gl/ptYCo>
- [33] Rendezvous Point Engineering. Cisco.com [online]. 2008 [cit. 2013-05-10]. Dostupné z: <http://goo.gl/ie497>
- [34] Cisco IOS IPv6 Configuration Guide: Release 12.2SX. In: Cisco IOS IPv6 Configuration Guide: Release 12.2SX [online]. 1. vyd.: CISCO, 2010 [cit. 2013-04-04]. Dostupné z: <http://goo.gl/FRGMc>
- [35] P. SAVOLA a B. HABERMAN. RFC 3956: Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address. In: Ietf.org: Network Working Group [online]. JHU APL, 2004 [cit. 2013-03-26]. Dostupné z: <http://www.ietf.org/rfc/rfc3956.txt>
- [36] Network Address Translation - NAT. BAČA, Radim. *Cs.usb.cz* [online]. ? [cit. 2013-05-10]. Dostupné z: <http://goo.gl/ie497>
- [37] Performance-Comparison Testing of IPv4 and IPv6 Throughput and Latency on Key Cisco Router Platforms: A Summary of Findings. In: Performance-Comparison Testing of IPv4 and IPv6 Throughput and Latency on Key Cisco Router Platforms: A Summary of Findings [online]. CISCO, 2007 [cit. 2013-04-04]. Dostupné z: <http://goo.gl/4cf3a>

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

BGP	Border Gateway Protocol
BSR	Bootstrap router
CPE	Customer-premises equipment
DHCP	Dynamic Host Configuration Protocol
DVMRP	Distance Vector Multicast Routing Protocol
EIGRP	Enhanced Interior Gateway Routing Protocol
IANA	Internet Assigned Numbers Authority
IGMP	Internet Group Management Protocol
IPTV	Internet Protocol television
IPv6	Internet Protocol Version 6
MLD	Multicast Listener Discovery
MOSPF	Multicast Open Shortest Path First
NIC	Network Interface Controller
OSPF	Open Shortest Path First
PIM	Protocol Independent Multicast
RFC	Request for Comments
RIP	Routing Information Protocol
RIPE	Réseaux IP Européens
RP	Rendezvous Point
RPF	Reverse Path Forwarding
RPT	Rendezvous Point Tree
SDP	Session Description Protocol
SPT	Source Path Tree
SSM	Source Specific Multicast
TCP	Transmission Control Protocol

SEZNAM OBRÁZKŮ

OBR. 1 - MULTICAST	4
OBR. 2 - MULTICAST ISO/OSI.....	4
OBR. 3 - STATICKÉ MAPOVÁNÍ MULTICASTOVÝCH MAC ADRES [1]	5
OBR. 4 – BITOVÝ ZÁPIS MULTICAST MAC ADRESY [2]	5
OBR. 5 - SHORTEST PATH TREES [4].....	9
OBR. 6 – SHARED DISTRIBUTION TREES [4]	10
OBR. 7 - PRVNÍ FÁZE PIM-SM – <i>RP TREE</i>	15
OBR. 8 - DRUHÁ FÁZE PIM-SM – <i>REGISTER-STOP</i>	16
OBR. 9 - TŘETÍ FÁZE PIM-SM – <i>SHORTEST-PATH TREE</i>	17
OBR. 10 – RP ELECTION PROCESS	18
OBR. 11 - PIM-SSM [10].....	19
OBR. 12 - ZACHYCENÁ ZPRÁVA IP <i>IGMP QUERY</i> A <i>IGMP REPORT</i>	22
OBR. 13 - IP DATAGRAM S IGMP ZPRÁVOU [13].....	23
OBR. 14 - OBECNÁ HLAVIČKA IP DATAGRAMU [14]	23
OBR. 15 - IGMP SNOOPING	24
OBR. 16 - GRAFICKÉ ZNÁZORNĚNÍ ŠKÁLOVÁNÍ MULTICASTU VE SROVNÁNÍ S UNICASTEM [15]	25
OBR. 17 – STATISTIKA PŘIJATÝCH PAKETŮ NA PORTU SWITCHE.....	26
OBR. 18 – POMĚR PAKETŮ NA INTERNETU	26
OBR. 19 – UNICAST	27
OBR. 20 – UNICASTOVÝ TCP THREE-WAY HANDSHAKE.....	28
OBR. 21 – BROADCAST	29
OBR. 22 – ZACHYCENÍ <i>ARP REQUEST</i> A <i>ARP REPLY</i>	30
OBR. 23 – ANYCAST	31
OBR. 24 – GEOCAST [18]	32
OBR. 25 - PŘÍJEM ZPOŽDĚNÝCH PAKETŮ [19]	34
OBR. 26 - OBRAZOVÉ ARTEFAKTY PŘI ZTRÁTĚ PAKETU.....	37
OBR. 27 - DSCP BITY V HLAVIČCE IP PAKETU [21]	38
OBR. 28 - 802.1Q A CoS V ETHERNETOVÉM RÁMCI [22].....	39
OBR. 29 - OBECNÁ TOPOLOGIE SÍTĚ.....	39
OBR. 30 - MOTOROLA VIP1003	40
OBR. 31 – ZJEDNODUŠENÁ TOPOLOGIE SÍTĚ	41
OBR. 322 - VÝPIS PIM-NEIGHBORS NA HRANIČNÍM ROUTERU	42
OBR. 33 - VÝPIS PIM INTERFACE NA VLAN 210 HRANIČNÍHO ROUTERU	43

OBR. 34 - ZOBRAZENÍ MULTICAST REVERSE PATCH FORWARDING KE ZDROJI	43
OBR. 35 - ALOKACE 172.17.5.216/29 NA VLAN 210	43
OBR. 36 - ALOKACE 172.17.240.0/23 NA VLAN 1559	44
OBR. 37 - VÝPIS IGMP SNOOPINGU	44
OBR. 38 - CPE JEDNOTKA	45
OBR. 39 - VÝCHOZÍ BRÁNY PRO IPTV	45
OBR. 40 - DISPLAY VERSION PIM ROUTERU	46
OBR. 41 - HP 5800-24G	46
OBR. 42 - KONFIGURACE HRANIČNÍHO PIM ROUTERU POSKYTOVATELE	47
OBR. 43 - PŘÍSTUPOVÝ SWITCH EDGE-CORE ES3528M	48
OBR. 44 - SHOW VERSION ACCESS SWITCHE	48
OBR. 45 - MULTICAST KONFIGURACE ACCESS SWITCHE	49
OBR. 46 - VÝPIS SHOW IP IGMP SNOOPING	49
OBR. 47 - MULTICAT	50
OBR. 49 - POUŽITÍ NÁSTROJE MULTICAT	50
OBR. 48 - STATICKÉ NASMĚROVÁNÍ SKUPINOVÉ ADRESY	50
OBR. 50 - TCPDUMP ODEBÍRANÉ SKUPINY	50
OBR. 51 - PERIODICKÉ ZASÍLÁNÍ HELLO MESSAGES	51
OBR. 52 - ZACHYCNÉ HELLO ZPRÁVY PIM ROUTERŮ NA SPOLEČNÉ DOMÉNĚ	52
OBR. 53 - POUŽITÍ JOIN/PRUNE ZPRÁV V PIM-DM	53
OBR. 54 - STATISTIKA JOIN/PRUNE/GRAFT ZPRÁVNA HRANIČNÍM ROUTERU	54
OBR. 55 - ZÁZNAM SKUPINY V PIM ROUTOVACÍ TABULCE	55
OBR. 56 - ZÁZNAM SKUPINY V OBECNÉ ROUTOVACÍ TABULCE	55
OBR. 57 - ZÁZNAM SAP SKUPINY V OBECNÉ ROUTOVACÍ TABULCE	55
OBR. 58 - NASTAVENÍ IGMP PARAMETRŮ KLIENTSKÉ IPTV VLAN	56
OBR. 59 - STATISTIKA IGMP ODBĚRATELŮ MULTICASTOVÉ SKUPINY	57
OBR. 60 - ZACHYCNÍ IGMP KOMUNIKACE SET TOP BOXU	57
OBR. 61 - PŘIDĚLENÍ IP SET TOP BOXU SLUŽBOU DHCP VZDÁLENÝM SERVEREM	58
OBR. 62 - MULTICASTOVÉ IPTV DATAGRAMY	59
OBR. 63 - MULTICASTOVÉ SDP DATAGRAMY	59
OBR. 64 - PŘEHLED ODEBÍRANÝCH MULTICASTOVÝCH SKUPIN NA PŘÍSTUPOVÉM SWITCHI	60
OBR. 65 - ZACHYCNÍ SPECIFIC IGMP QUERY	60
OBR. 66 - UNICASTOVÝ PŘENOS NELINEÁRNÍHO OBSAHU	61
OBR. 67 - TEORETICKÉ ŠKÁLOVÁNÍ MULTICASTU A UNICASTU	62
OBR. 68 - IPV6 MULTICASTOVÁ ADRESA [24]	63
OBR. 69 - MAPOVÁNÍ IPV6 MULTICAST ADRESY NA MAC ADRESU [27]	66

OBR. 70 – FORMÁT MLDV1 A MLDV2 [30]	68
OBR. 71 - VTISKNUTÍ <i>RENDEZVOUS POINT</i> DO IPV6 ADRESY (<i>RFC 3956</i>) [33]	70
OBR. 72 - VÝPIS <i>SHOW VERSION</i> POUŽITÉHO ROUTERU CISCO 2691.....	71
OBR. 73 - SÍŤOVÁ TOPOLOGIE PRO SIMULACI SMĚROVÁNÍ IPV6 MULTICASTŮ.....	72
OBR. 74 - SPECIFICKÉ NASTAVENÍ RP A C1 (C2)	73
OBR. 75 - MOŽNOSTI ANALÝZY IPV6 SYSTÉMU IOS.....	74
OBR. 76 - VYTVOŘENÝ SPT DISTRIBUČNÍ STROM NA DR MULTICASTOVÉHO ZDROJE	75
OBR. 77 - GENEROVÁNÍ MULTICASTOVÉHO PROVOZU PROSTŘEDNICTVÍM ICMPV3 ZPRÁV	75
OBR. 78 - ROUTER INFORMUJE O NOVÉM ZDROJI A SMĚRUJE PROVOZ NA RP.....	75
OBR. 79 - POSLUCHAČI V MLD TABULCE ROUTERU R1.....	76
OBR. 80 - VYTVOŘENÉ DISTRIBUČNÍ STROMY NA R1 (DR PRO C1).....	76
OBR. 81 – PAT A SNAT [36]	78
OBR. 82 – SROVNÁNÍ NÁROČNOSTI SMĚROVÁNÍ IPV4 A IPV6	79
OBR. 84 – IPV4 PŘENOSY NIX.CZ (DO 21.3.2013)	80
OBR. 84 – IPV6 PŘENOSY NIX.CZ (DO 21.3.2013)	80

SEZNAM TABULEK

TAB. 1 - MAX. DOPORUČENÉ HODNOTY MLR	35
TAB. 2 - DOPORUČENÍ CoS HODNOT	39
TAB. 3 - PŘEHLED ADRESOVÁNÍ IPV6 MULTICASTU	64
TAB. 4 - <i>WELL-KNOWN IPV6 MULTICAST ADDRESSES</i>	65
TAB. 5 - ZPRÁVY MULTICAST LISTENER DISCOVERY	67
TAB. 6 - NASTAVENÍ JEDNOTLIVÝCH SÍŤOVÝCH PRVKŮ A JEJICH ROZHRANÍ	72

SEZNAM PŘÍLOH

1. P1 – TRACEROUTE ANYCASTOVÉ ADRESY 8.8.8.8
2. P2 – ELEKTRONICKÁ PŘÍLOHA – KONFIGURACE SÍŤOVÝCH PRVKŮ
3. P3 – DATASHEETY

PŘÍLOHA P I: TRACEROUTE ANYCASTOVÉ ADRESY 8.8.8.8

```
Czech republic:~# traceroute -n 8.8.8.8
traceroute to 8.8.8.8 (8.8.8.8), 30 hops max, 60 byte packets
 1  93.89.96.1           1.592 ms   1.643 ms   1.649 ms
 2  81.201.48.195        8.472 ms   8.512 ms   8.415 ms
 3  81.201.48.194        8.489 ms   8.741 ms   8.678 ms
 4  72.14.220.74         8.010 ms   7.981 ms   7.986 ms
 5  216.239.46.141      15.697 ms  24.147 ms  24.043 ms
 6  72.14.239.60         16.287 ms
    72.14.236.68        16.274 ms   16.210 ms
 7  209.85.254.116      16.096 ms
    209.85.254.114      15.988 ms
    209.85.254.116      16.036 ms
 8  * * *
 9  8.8.8.8              16.078 ms  16.037 ms  16.074 ms
```

```
Telstra.net, Austrálie @ telstra.net/cgi-bin/trace
```

```
 1  203.50.77.49         0.306 ms   0.281 ms   0.244 ms
 2  203.50.80.1          2.992 ms   2.049 ms   3.140 ms
 3  203.50.11.74         17.715 ms  15.416 ms  15.986 ms
 4  203.50.6.154         21.230 ms  23.161 ms  23.980 ms
 5  74.125.50.1          15.236 ms  15.292 ms  15.235 ms
 6  66.249.95.224        15.486 ms  15.542 ms
    66.249.95.226        15.486 ms
 7  66.249.95.233        53.645 ms
    66.249.95.235        15.418 ms   18.416 ms
 8  72.14.237.21         15.609 ms  15.663 ms  15.610 ms
 9  8.8.8.8              15.360 ms  15.418 ms  15.360 ms
```

```
University of Southern California, USA @ usc.edu/cgi-bin/traceroute
```

```
 3  68.181.194.65        1.094 ms   0.709 ms   0.634 ms
 4  137.164.23.225       1.038 ms   0.919 ms   0.856 ms
 5  137.164.46.132       3.228 ms   3.338 ms   3.690 ms
 6  72.14.223.85         1.914 ms   2.435 ms   1.265 ms
 7  216.239.46.40        4.100 ms   2.116 ms
    64.233.174.238       1.021 ms
 8  72.14.238.2          28.895 ms
    72.14.238.0          28.421 ms   28.225 ms
 9  72.14.239.162        38.097 ms  27.520 ms  27.192 ms
10  64.233.174.131       28.561 ms
    64.233.174.129       28.434 ms   28.080 ms
11  * * *
12  8.8.8.8              27.412 ms  27.447 ms  27.309 ms
```

PŘÍLOHA P II:

EL. PŘÍLOHA - KONFIGURACE SÍŤOVÝCH PRVKŮ

1. IGMP – CPE
2. IGMP - access_switch
3. IGMP - aggregation_switch
4. PIM – router
5. IPv6 – multicast

PŘÍLOHA P III:

EL. PŘÍLOHA - DATASHEETY

- | | |
|------------------------------|-------------------|
| 1. Hraniční router/L3 switch | HP A5800-24 |
| 2. AgregáčnÍ switch | DCN DCRS-5750F |
| 3. Přístupový switch | Edge-Core es3528m |
| 4. Set top box | Motorola VIP1003 |