IPTV protokollok

oktatási segédanyag az „IP alapú távközlés” c. tárgyhoz

Készült:

Steierlein Balázs:

IPTV rendszerek

című szakdolgozatának felhasználásával.

Szerkesztette: Lencse Gábor

Az anyag témakörei:

- IP multicast fogalma, IP (v4, v6) multicast címzése, Ethernet multicast címek származtatása.
- IGMP, IGMP snooping, MLD.
- Multicast routing protokollok közül: PIM-SM, PIM-DM

Az apró betűvel szedett részek tájékoztatásul szolgálnak, ezeket nem kell megtanulni!

2011. 10. 14.

Tartalom

1. IP Multicast

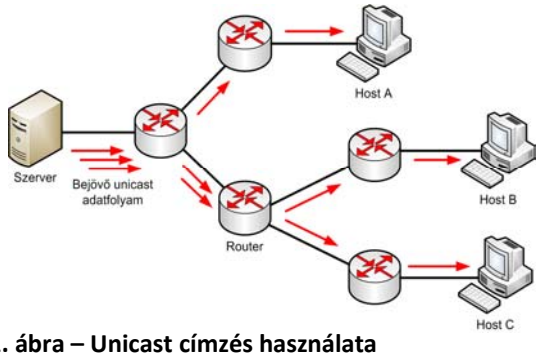
RFC 1112

Hagyományos IP hálózatokban a csomagokat a forrás vagy *unicast* vagy *broadcast* címzéssel küldi el, ami bizonyos multimédiás alkalmazások/szolgáltatások esetében nem hatékony sávszélesség kihasználást eredményez. Tipikusan ilyen szolgáltatás az IPTV.

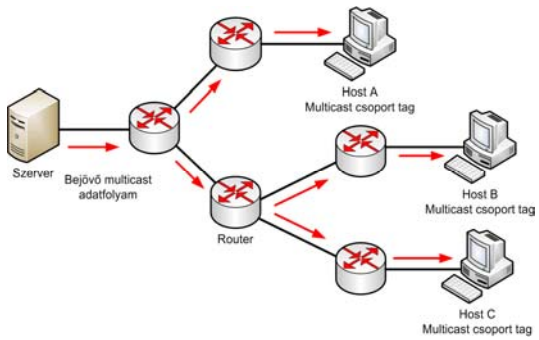
Unicast címzés esetén a forrásnak ugyanazt a csomagot az összes felhasználónak külön meg kell címeznie, és elküldenie, ami a hálózatnak a forráshoz közeli részén komoly sávszélesség igényével járna.

Ha pedig broadcast címzéssel szeretnénk megoldani a feladatot, akkor a hálózaton lévő összes felhasználó megkapná az adatokat, azok is, akik nem is tartanak rá igényt, ezzel fölöslegesen terhelve a hálózatot.

Ezért *multicast* címzést alkalmaznak. A multicast IP címek csoportokat határoznak meg, amikről a hostok eldönthetik, hogy akarnak-e a csoport tagjai lenni vagy sem. A forrás a *csoportcím*re (és ezáltal magának a csoportnak) küldi az adatokat, és így mindazok a hostok megkapják a streamet, akik éppen akkor a csoportban vannak. A csoporton kívüli hostokhoz nem jut el a stream, vagyis fölös adatforgalom nincs. Az unicast, és a multicast adatforgalom közti különbséget szemlélteti az alábbi két ábra.



1. ábra – Unicast címzés használata



2. ábra – Multicast címzés használata

1.1 MULTICAST HÁLÓZATI CÍMZÉS IPV4-ES ÉS IPV6-OS KÖRNYEZETBEN

IPv4-es környezetben multicast címzésre a D osztályú IP címek (vagy hívják még GDA-nak: Group Destination Address [37]) vannak fenntartva, ami a 224.0.0.0-239.255.255.255 címtartományt jelenti. Ezen címek első négy bitje az 1110 prefix. Fontos megjegyezni, hogy ezek a multicast csoportoknak a címei, nem pedig a forrásé, vagy akár a klienseké, nekik ugyanúgy unicast IP címük van.

Két féle multicast csoportot különböztetünk meg: *állandó csoport* és *tranziens csoport*.

Az **állandó csoport** akkor sem szűnik meg, ha az utolsó tag is kilépett. Ezek a csoportok általában a routing protokolloknak, és egyéb topológia felderítő mechanizmusoknak vannak fenntartva. A számunkra fontosabbakat az alábbi listában soroljuk fel, a három legfontosabb (=megtanulandó) csoportot vastag betűvel emeljük ki [30][35]:

224.0.0.1	**Az összes host címzése a hálózatban**
224.0.0.2	**Az összes router címzése a hálózatban**
224.0.0.4	Az összes DVMRP router címzése a hálózatban
224.0.0.5	Az összes OSPF router címzése a hálózatban
224.0.0.13	**Az összes PIM router címzése a hálózatban**
224.0.0.15	Az összes CBT router címzése a hálózatban
239.0.0.0/24	Privát tartomány

A **tranziens csoportok** rendszerint valamely feladat idejére jönnek létre, és ha kilépett az utolsó tag is, akkor megszűnnek.

A 224.0.0.0/24 címtartomány úgynevezett link-local tartomány, az ide tartozó címekre küldött üzeneteket a routerek nem továbbítják.

A privát tartományt pedig bárki felhasználhatja a saját rendszerén belül (a unicast privát IP címekhez hasonlóan.)

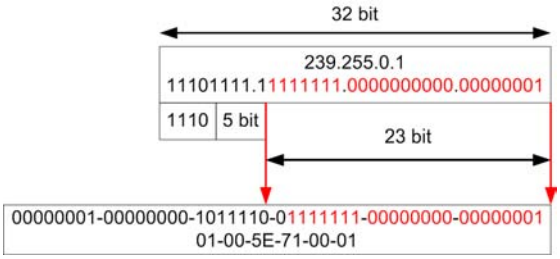
IPv6-os környezetben az FF00::/8 prefix-szel kezdődnek a multicast címek. Természetesen itt is vannak fenntartott csoportcímek [35][36]:

FF02::2	Az összes router címzésére való
FF02::4	Az összes DVMRP router címzésére való
FF02::5	Az összes OSPF router címzésére való
FF02::D	Az összes PIM router címzésére való
FF02::16	Az összes MLDv2 router címzésére való

1.2 MULTICAST ETHERNET CÍMZÉS IPV4-ES ÉS IPV6-OS KÖRNYEZETBEN

RFC 1112, 2710

A harmadik rétegbeli IP címeket le kell képeznünk a második rétegbeli MAC csoport címekre, ha Ethernet hálózatokon akarunk adatokat átvinni. Logikus gondolat, hogy az éppen használt multicast IP-címből képezzük a MAC csoportcímet.



3. ábra – Multicast MAC címek képzése multicast IP címekből

A leképzés megértéséhez vizsgáljuk meg a MAC címek felépítését! Először is az OUI (Organizationally Unique Identifier) első[1] bitje határozza meg, hogy a kérdéses MAC cím az multicast cím-e (ebben az esetben ez a bit 1), vagy egy hálózati kártya egyedi címéről van-e szó (ekkor a bit 0). Az IPv4 címből képzett multicast MAC címek első 25 bitje előre meg van határozva. A multicast MAC címek képzésekor a 01:00:5E:00:00:00 Ethernet multicast cím felső 25 bitjét megtartjuk és az alsó 23 bitjébe illesztjük bele az IPv4 cím alsó 23 bitjét. Ennek a megoldásnak hátránya, hogy mivel nem a teljes IP cím kerül leképezésre, így több multicast IP címnek lesz ugyanaz a MAC címe. (Mivel minden multicast IP cím az 1110 prefix-szel kezdődik, így összesen 5 bitet veszítünk el.) [16]

IPv6-nál egyszerűbb a helyzet. A MAC cím első két bájtja fixen a hexadecimális 33:33 lesz, a többi pedig az IPv6 cím alsó 4 bájtja. [31]

[1] A MAC címek *lsb* (least significant bit first = legkisebb helyiértékű bit elöl) bitsorrendben kerülnek továbbításra, így itt a közegen legelsőként továbbított bitről, azaz a cím első oktetjének legkisebb helyiértékű bitjéről van szó.

2. Csoportmenedzsment

2.1 Internet Group Management Protocol (IGMP)

RFC 2236, RFC 3376

Az IGMP protokollt a hostok multicast csoporthoz való csatlakozásukhoz, illetve onnan való kilépésükhöz használják. Ez a hostok és a first-hop router között történik. Három verziója van, amiből az IGMPv2 a legáltalánosabban használt az IPTV rendszerekben, de az IGMPv3 sok olyan újabb funkcióval rendelkezik, amik miatt később átveheti az IGMPv2 helyét a gyakorlatban. (A magasabb verziószámú protokollok természetesen visszafele kompatibilisek.)

Az IGMPv2 és az IGMPv1 is az *Any Source Multicast* (ASM) hálózatokat támogatják, ami azt jelenti, hogy a host ugyan megválaszthatja, hogy melyik csoporthoz akar csatlakozni, de a csoporton belül minden forgalmat megkap, függetlenül attól, hogy ki a forrás.

2.1.1. Egy host csoporttagsága

Csoporttagság lekérdezése

A *multicast router* periodikusan küld egy IGMP ***Host Membership Query*** üzenetet (röviden query), amivel megtudakolja, hogy mely csoportoknak vannak tagjai. Ezt úgy oldja meg, hogy ezt az IGMP query-t az all-hosts IP multicast címre küldi (224.0.0.1). Ezt mindenki megkapja, aki az adott hálózaton tartózkodik. A query-t tartalmazó IP datagramban lévő TTL mező értéke 1. Így biztosítja, hogy a query csak a hostokig menjen. A csoport tagjai erre válaszolnak egy ***Host Membership Report*** üzenettel (röviden report). A report üzenetek torlódásának elkerülése érdekében nem mindenki egyszerre küldi a report üzenetet, hanem csak egy random timer lejárta után. A hostok a csoport címére címezik a report üzenetet, így mindenki megkapja a csoportban. Így ha egy host a timerének lejárta előtt megkapja egy másik host által küldött report üzenetet, akkor visszavonja a sajátját. Ezáltal csak egy report üzenetet kap meg a router. De elég egyetlen report üzenet is, mivel ha a csoportnak csak egyetlen tagja van, a routernek akkor is küldenie kell multicast streamet.

Mivel a multicast router periodikusan küldi a query üzeneteket, így állandóan frissítve vannak az információi arról, hogy van-e még tagja a csoportnak.

Belépés egy csoportba

Ha a host csatlakozik egy új, még üres csoporthoz, akkor nem várja meg míg queryt kap, hanem rögtön küld egy report üzenetet a routernek. Ezzel jelezvén, hogy ő az első tagja a csoportnak. Annak érdekében, hogy ez a report nehogy elvesszen, a hostnak javasolt rövid időn belül megismételnie ezt az üzenetet.

Csoport elhagyása

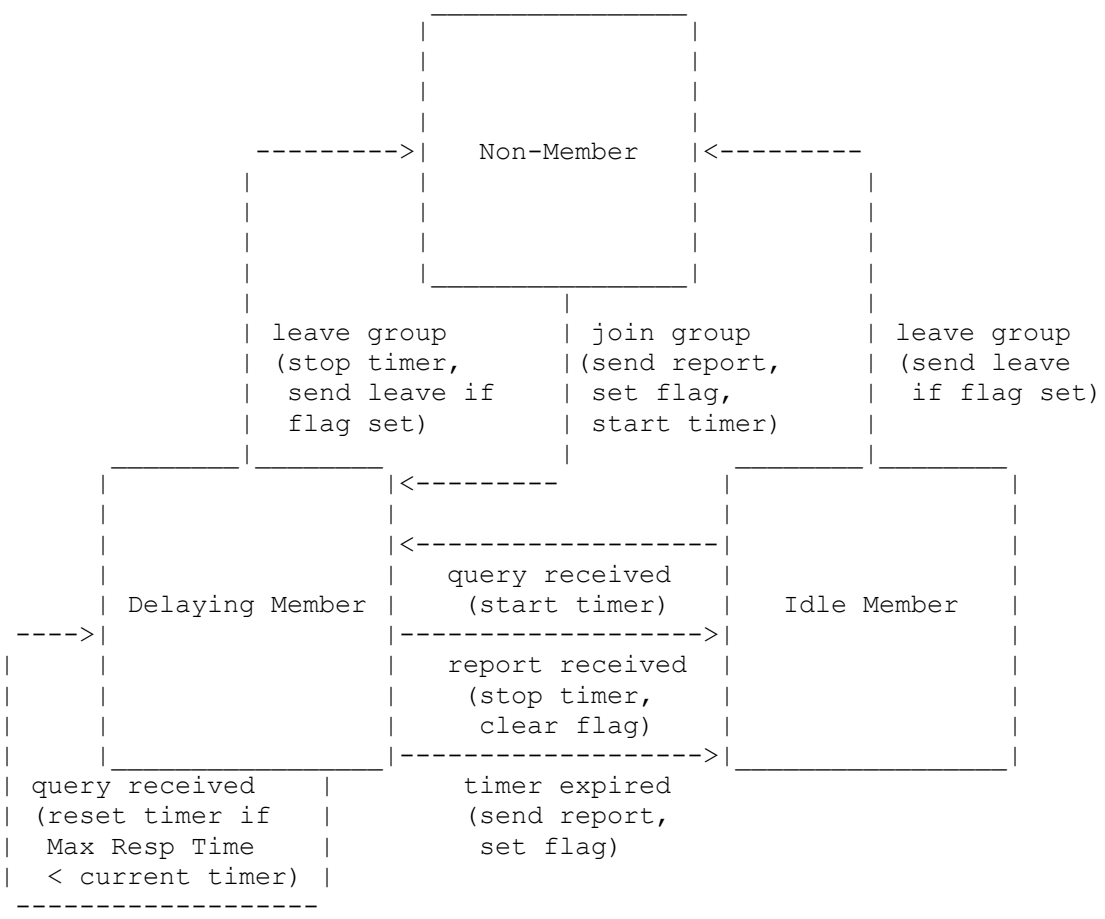
A fontos újítás volt az egyes verzióhoz képest, hogy a host képes ***Leave*** üzenetet küldeni. (Lényegében e funkció hiánya miatt nem volt alkalmas az IGMPv1 az IPTV rendszerekben a csoportmenedzsmentre.) Ezt az üzenetet az all-routers multicast címre küldi (224.0.0.2). Amikor egy multicast router router kap egy ilyen Leave üzenetet, akkor utána küld egy ***csoport-specifikus query*** üzenetet, ami, ahogy a nevében is benne van, csak annak a csoportnak szól, ahonnan a Leave üzenet érkezett. Ha erre nem válaszol senki a maximális válaszadási idő (***Max Response Time***) lejárta előtt, akkor a router tudomásul veszi, hogy üres a csoport. Ezzel csökken annak az ideje, amíg kiderül a router számára, hogy nincs hallgatója a streamnek. Míg az IGMPv1 esetében a legrosszabb esetben a random timer (max 10 másodperc) + query periódus (60 másodperc) ideig tartott, amíg kiderült a router számára, hogy üres a csoport. Tehát akár 70 másodpercig is áramolhatott a stream úgy, hogy nem is volt vevő a csoportban, és ezzel fölöslegesen terhelődött a hálózat.

Egy host állapotátmenet diagramja

Egy host három állapotban lehet:

- Nem tag: amikor nem tagja semmilyen csoportnak.
- Késleltetett tag: amikor tagja a csoportnak és fut a timer.
- Tétlen tag: szintén tagja egy csoportnak, de nem fut a timer.

A host egy jelzőbitet (flag) használ annak nyomon követésére, hogy egyedül van-e a csoportjában (flag=1), vagy más tagja is van a csoportnak (flag=0). A Leave üzenetet csak az előbbi esetben kell ténylegesen elküldenie. Az alábbi ábrán követhetjük nyomon egy host pontos viselkedését.



4. ábra – Egy host állapotátmenet diagramja

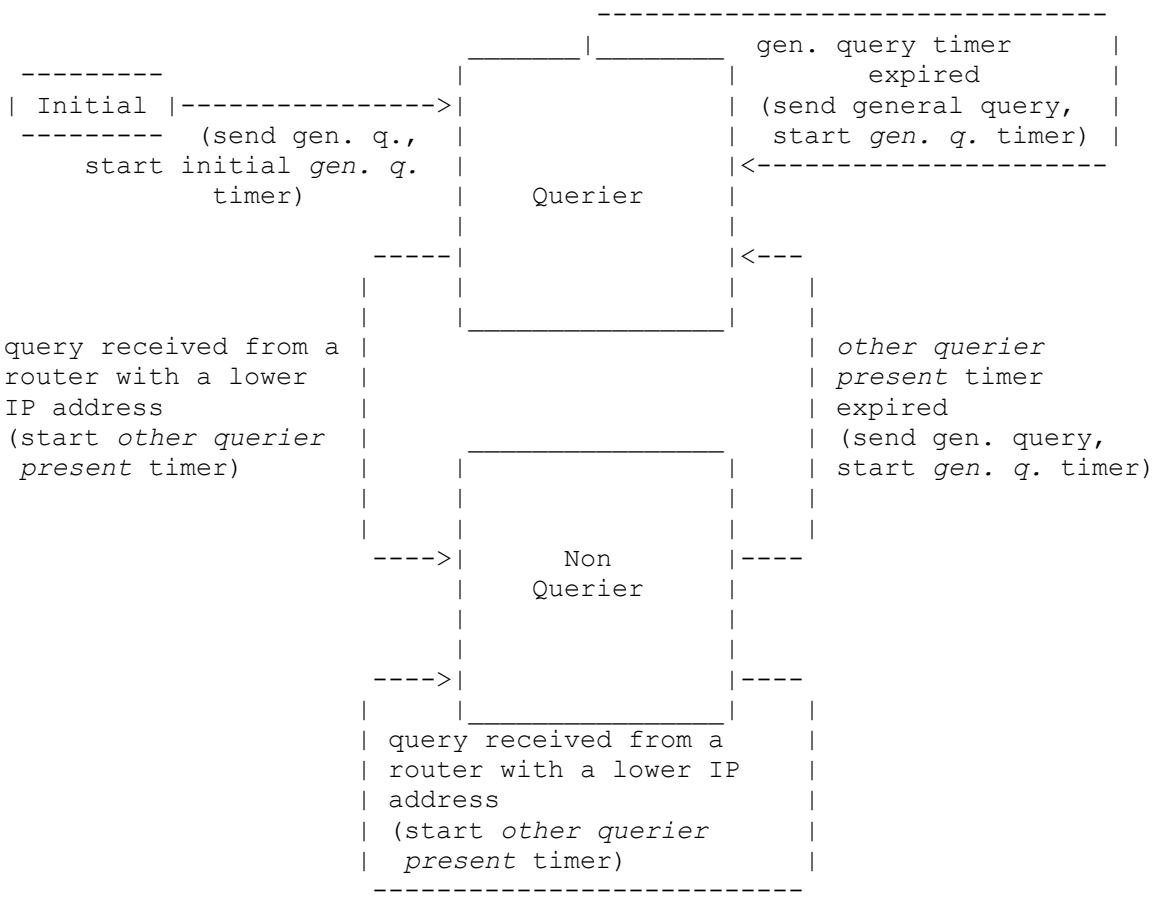
2.1.2. IGMP Querier Election

Ha több IGMP képes router is van egy szegmensen, akkor az ***IGMP Querier Election*** mechanizmust használják annak eldöntésére, hogy ki legyen az, aki majd a queryket fogja küldeni a hálózaton. Ez egész egyszerűen úgy működik, hogy mindig a legalacsonyabb IP című router lesz a Querier. Ennek kiderítésére a routerek periodikusan küldenek általános queryket minden hozzájuk kapcsolódó hálózatra. Minden IGMP képes router alapból querierként kezd, és csak utána lépnek vissza non-querier állapotba.

Egy router állapot-átmenet diagramja

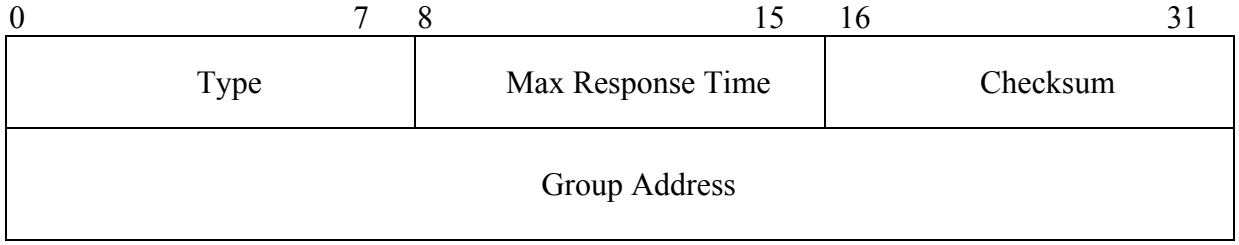
Egy router két állapotban lehet:

- Querier: ez a router van kijelölve a query üzenetek küldésére
- Non-Querier: ez a router nem küld query üzeneteket



5. ábra – Egy router állapotátmenet diagramja

2.1.3 Az IGMP üzenetek felépítése



6. ábra – Egy IGMPv2 üzenet felépítése

Type: az üzenet típusa. Háromféle üzenettípus van: Membership Query, Membership Report, Leave Group. A Membership Querynek pedig két fajtája van:

- General Query: arra való, hogy a megtudja a router, hogy mely csoportnak vannak még tagjai: a csoportcím mező értéke: 0.0.0.0
- Group-Specific Query: ezzel egy konkrét csoportnál érdeklődünk, vannak-e tagjai.

Max Response Time: maximális válaszadási idő, tized másodpercekben mérik. Az a maximálisan megengedett idő, amíg a Query üzenetre a Report üzenetnek meg kell érkeznie.

Mivel csak a Query üzenetek szempontjából van jelentősége, ezért a többi üzenettípusnál az értéke nullára van beállítva, és a címzettek figyelmen kívül hagyják.

Checksum: ellenőrző összeg.

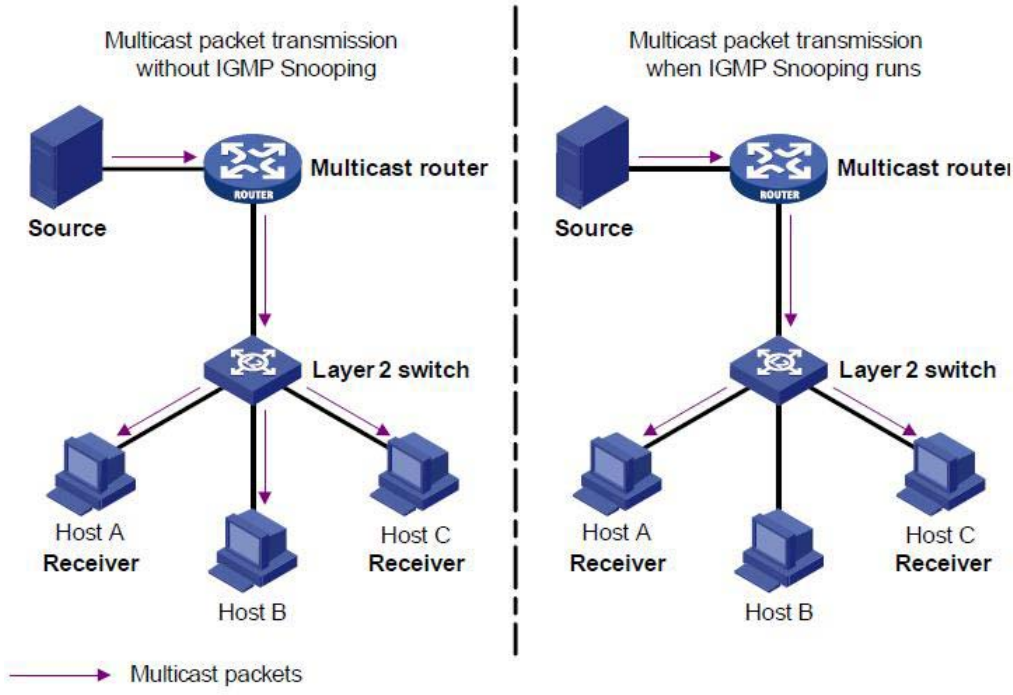
Group Address: Membership Query üzeneteknél, ha az üzenet General Query, akkor az értéke nulla, ha Group-Specific, akkor a csoport címe szerepel itt. Report vagy Leave üzeneteknél annak a csoportnak az IP multicast címét tartalmazza, ahova vagy ahonnan a host be- vagy kilépett.

2.1.3 IGMP Snooping

Az IGMP snooping feladata

Az IGMP Snooping funkció arra szolgál, hogy a multicast router és a hostok között lévő switchek a multicast forgalmat csak azon portjaikra küldjék ki, ahol igény van a forgalom vételére. IGMP Snooping nélkül a switch broadcast szerűen árasztaná el adatokkal a hozzá kapcsolódó hostokat.

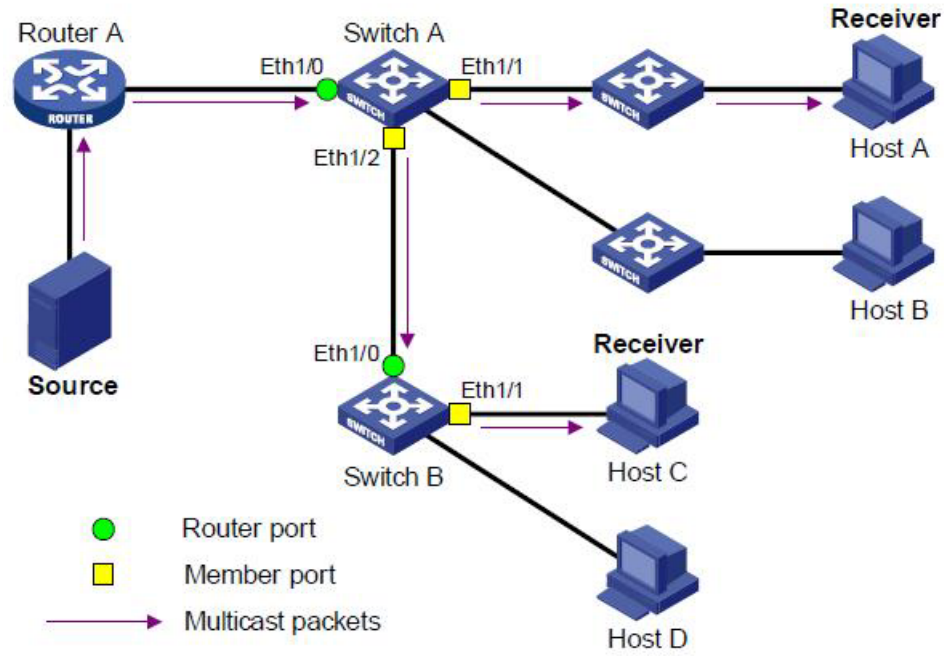
Vegyük észre, hogy bár a switch egy adatkapcsolati rétegben működő eszköz, az IGMP snooping funkció egy magasabb rétegbeli protokoll mezőinek vizsgálatát, sőt a protokollban való részvételét teszi szükségessé!



7. ábra – Az IGMP Snooping jelentősége

Elnevezések

A könnyebb érthetőség és egyszerűség érdekében definiáljunk két fogalmat a switchekkel kapcsolatban: ***router port***, és ***member port***. A router port jelentse a switchnek azt a portját, ami a router irányába „mutat", a member port pedig azt, amin keresztülhaladva egy csoport tag felé közelednénk. A switchnek tisztában kell lennie azzal, hogy melyik portján vannak routerek, és melyeken kliensek, azaz csoporttagok.



8. ábra – A *router port* és a *member port* illuszrációja

Működés

A routerek megtalálásának érdekében a switch következő üzeneteket figyeli (zárójelben az Ethernet multicast cím szerepel):

- **IGMP Membership Query (01-00-5E-00-00-01)**
- **PIMv2 Hello üzenet (01-00-5E-00-00-0D)**
- DVMRP Probe üzenet (01-00-5E-00-00-04)
- MOSPF üzenet (01-00-5E-00-00-05)

(Az utóbbi kettővel mi most nem foglalkozunk.)

Ha a switch észlel egy routert valamelyik portján, akkor az a port bekerül a router portok listájába.

Ha egy kliens elsőként csatlakozik egy multicast csoporthoz, akkor az IGMP Membership Report üzenetének hatására a switch létrehoz egy bejegyzést annak a csoportnak, amelyikhez a kliens csatlakozni szeretetne. A csoporthoz pedig hozzárendeli az összes router portot, és azt

a member portot, amelyiken a Report üzenet érkezett. Ezután továbbítja a Report üzenetet az összes router portjára.

Ha a kliens nem elsőként csatlakozik egy csoporthoz, akkor folyamat annyiban módosul, hogy a switch nem fogja továbbítani a kliens Report üzenetét, hanem csak csoportonként egyet, 10 másodperces periódusokban.

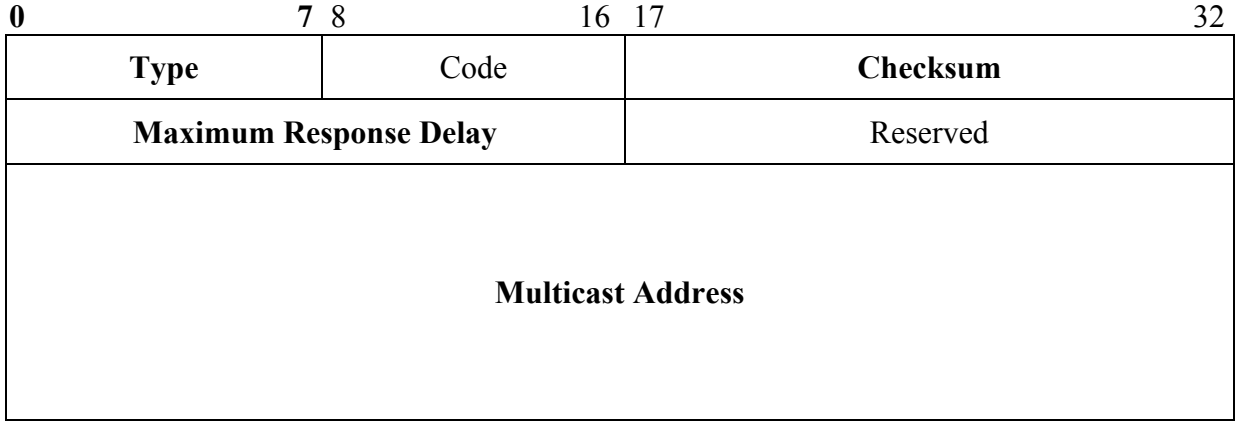
Egy tag kilépésekor a switch veszi a Leave üzenetet, és küld egy csoport-specifikus Queryt arra a portra (és csak arra) amelyiken a Leave üzenet érkezett. Ha nem érkezik válaszként Report üzenet, a switch törli a portot a bejegyzésből. Abban az esetben, ha a kilépő tag az utolsó volt, akkor a switch továbbítja a Leave üzenetet az összes router portjára, és törli a csoport bejegyzését. [37]

2.2 MULTICAST LISTENER DISCOVERY (MLD)

RFC 2710

Az MLD protokoll funkciója és működése gyakorlatilag megegyezik az IGMP-vel, csak IPv6-os környezetben. Az MLD az ICMPv6 üzenetformátumát használja üzenetei szállítására.

Az MLD üzenetformátuma (a vastag betűvel írt nevű mezőknek van IGMP megfelelője):



9. ábra – Egy MLD üzenet felépítése

Type. Az üzenetnek háromféle típusa lehet:

- Multicast Listener Query: két fajtája van: általános query és csoportcím-specifikus query. Ez elsővel megtudhatjuk, hogy mely csoportoknak vannak tagjai, a másodikkal pedig konkrét csoportról tudhatjuk meg, hogy vannak-e tagjai.
- Multicast Listener Report: a vevő ezzel jelzi igényét a multicast streamre. A Multicast Address mezőben ilyenkor az a csoportcím szerepel, amin hallgatni szeretne.

- Multicast Listener Done: lényegében ez a „Leave” üzenet, ha a vevő nem kívánja tovább venni a streamet.

Code. A forrás nullára állítja, a vevő figyelmen kívül hagyja.

Checksum. Ellenőrző összeg.

Maximum Response Delay. Maximum várakozási idő lényegében. Csak Query üzeneteknél értelmezzük, milliszekundumban. A többi üzenet fajtánál ennek a mezőnek az értékét a küldő nullára állítja, a vevő pedig nem veszi figyelembe.

Reserved. Fenntartott mező, a forrás nullára állítja, a vevő figyelmen kívül hagyja.

Multicast Address. Általános Query üzenetnél a mező értéke nulla. Report, Done, és csoportcím-specifikus Query üzeneteknél pedig az adott multicast csoportcím szerepel itt.

3. Multicast Routing Protokollok

Természetesen a multicast adatátvitelhez multicast routing protokollokat kell használnunk. Ilyenek például: PIM-SM, CBT, PIM-DM, DVMRP, MOSPF.

A ***Protocol Independent Multicast***, ahogy a nevében is benne van, független multicast protokoll, ami azt jelenti, hogy nincs saját topológia felderítő algoritmusa, hanem egy másik protokoll (pl. RIP, OSPF, BGP) adatait használja fel. (Az első verziót még 1995-ben alkották meg, de sosem szabványosította az IETF. A kettes verziót 1997-ben szabványosították, és 1998-ban újították meg.) A PIM úgy működik, hogy a külső topológia információkon alapuló útvonalakon épít ki multicast fákat a forrástól a vevőig. Alapvetően arra tervezték, hogy egy autonóm rendszeren (AS) belül működjön. Két fő fajtája létezik:

- ***PIM – Sparse Mode*** (ritka): ez a legszélesebb körben elterjedt multicast routing protokoll, ami nem feltételez mindenütt csoporttagokat, ezért csak arra küld multicast forgalmat, ahol arra igény van.
- ***PIM – Dense Mode*** (sűrű): ez ritkábban használt, mivel úgy építi ki a fákat, hogy elárasztja az egész hálózatot multicast forgalommal, és ahol nincsenek vevők, azokat az ágakat visszametszi.

3.1.1 Protocol Independent Multicast – Sparse Mode (PIM-SM)

A PIM-SM pontos leírása az RFC 2362-ben található.

Mint már fent említettük, a PIM-SM-nek nincs saját topológia felderítő mechanizmusa, ezért egy külső protokoll routing tábláját használja fel a sajátjának elkészítésére. Ezt a saját routing táblát hívjuk MRIB-nek (*Multicast Routing Information Base*).

A PIM-SM protokoll működésének 3 fázisa van.

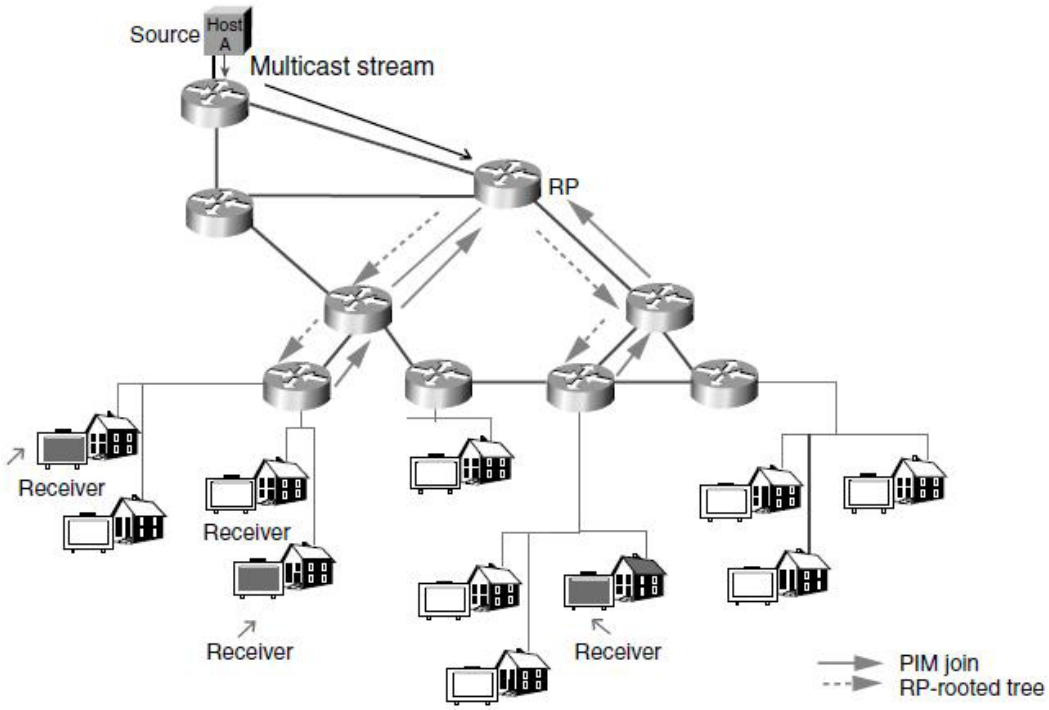
Első fázis: RP fa (RPT= *Rendezvous Point Tree*) kiépítése a vevők felől

Nevezik ezt még megosztott fának is, mert akár az összes forrás forgalma áthaladhat a fán.

Minden AS-ben szükség van egy ***Találkozási Pont***ra (*Rendezvous Point*, továbbiakban RP). Az RP egy PIM-SM router, ami a forrás és a vevők között segít kiépíteni a kapcsolatot. Az RP címét vagy statikusan beállítják az AS összes routerén, vagy egy megfelelő algoritmussal megválasztják az AS PIM-SM routerei közül (lásd az apró betűvel írt részben: RP információk vétele).

Továbbá a vevő helyi PIM routerei közül választanak egy ***Kijelölt router***t (*Designated Router*, továbbiakban DR). (A választás módjáról lásd az apró betűvel írt részben a Hello üzenet leírását.)

Az első fázisban a vevő jelzi igényét a multicast forgalomra, ami egy adott csoportnak van címezve. Ehhez IGMP, vagy MLD protokollt használ, attól függően, hogy IPv4, vagy IPv6-os környezetről van szó. Ha a DR megkapja a vevőtől a kérést, akkor küld egy ***PIM (*, G) Join*** üzenetet a kívánt multicast csoportra vonatkozóan az RP felé (az IP célcím: az RP unicast címe). A (*, G) jelölésben az első tag (a „*") a forrás, a második tag (a „G") a multicast csoport címe. A * azt jelenti, hogy a csoportba való csatlakozáskor az összes olyan forrás forgalmát megkapjuk, akik annak a csoportnak küldik a streamet. A Join üzenet végighalad a routereken keresztül az RP-ig, és kiépül a fa, aminek a gyökere az RP router. Valójában nem feltétlenül kell egészen az RP-ig haladnia a Join üzenetnek, hanem elég csak addig a routerig, ahol már a fa ki van épülve. A Join üzeneteket addig újraküldik, amíg a vevő a csoport tagja marad. Ha az összes hozzá kapcsolódó vevő elhagyja a csoportot, a DR küld egy PIM (*, G) Prune üzenetet, amivel a fa visszametsződik addig a routerig, amihez már kapcsolódnak forgalmat fogadó vevők.



10. ábra – Az RP fa kiépülése

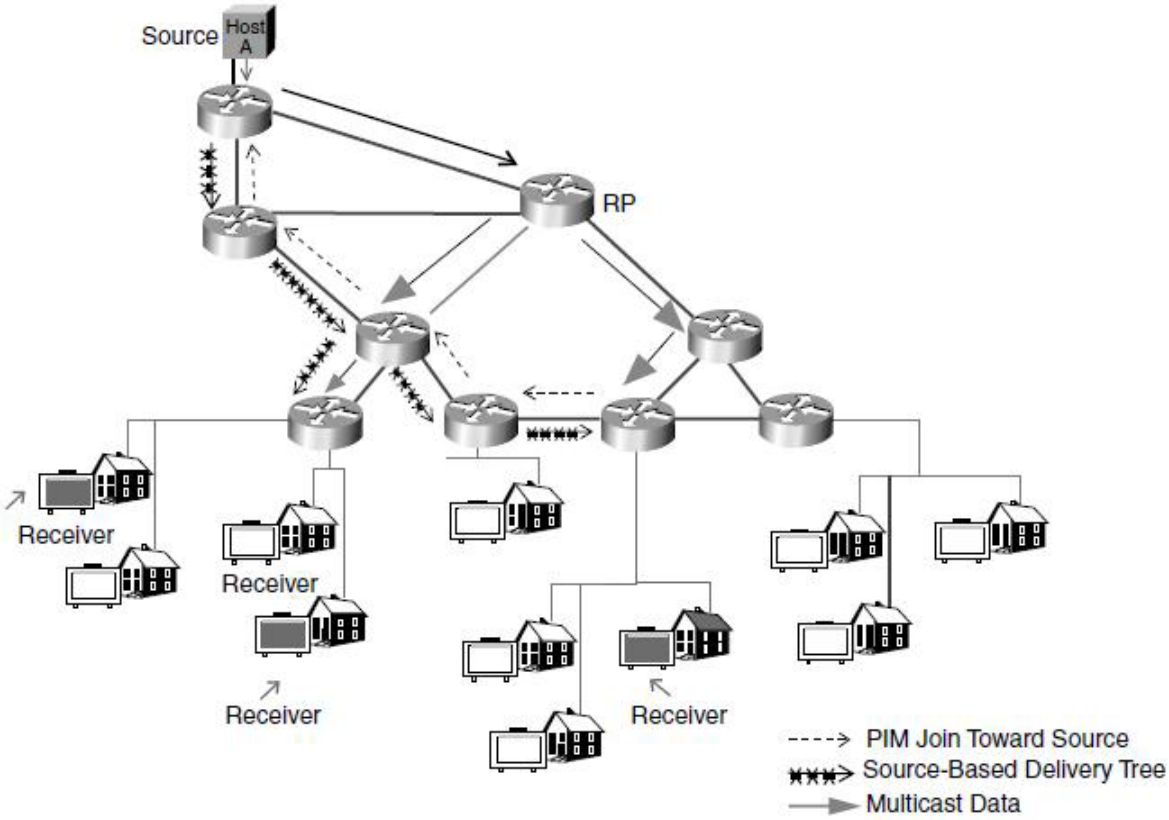
Második fázis: Regisztráció, SPT kiépítése az RP felől a forrás felé, Register-stop

A források első multicast csomagjait a first-hop router (DR) unicast IP csomagokba ágyazzák, és így küldik el az RP-nek. Ezeket a csomagokat hívjuk regiszter csomagoknak. A forrás ezzel az üzenettel tudatja az RP-vel, hogy készen áll adatfolyam küldésre. Az RP kicsomagolja, és megnézi, hogy melyik multicast csoport a címzett, majd továbbítja. A multicast így már működőképes, de korántsem optimális: az RP-nek kell kicsomagolnia és továbbítani az összes multicast csomagot (ez minden egyes csomagnál feldolgozást igényel, és az útvonal sem feltétlenül optimális). Ennek orvoslására először is az RP küld egy PIM (S, G) Join üzenetet a forrás felé. Az üzenet végighalad az útba eső routereken, és azok bejegyzik a táblájukba a (S, G) párost, ha még nem volt ilyenjük. Amikor a Join üzenet eléri az S forrás DR-ét vagy egy olyan routert, aminek már van (S, G) bejegyzése, akkor az adatok elkezdenek áramlani az S forrástól az RP felé, immár multicast módon. Ezzel kiépült az SPT (Shortest Path Tree) a forrás és az RP között. Ezt a folyamatot nevezzük a forrás regisztrációjának.[1] Ekkor az RP küld egy Register-Stop üzenetet, hogy most már nincs szükség rá, hogy a forrás DR-e unicast csomagokba ágyazza a multicast adatokat. [5]

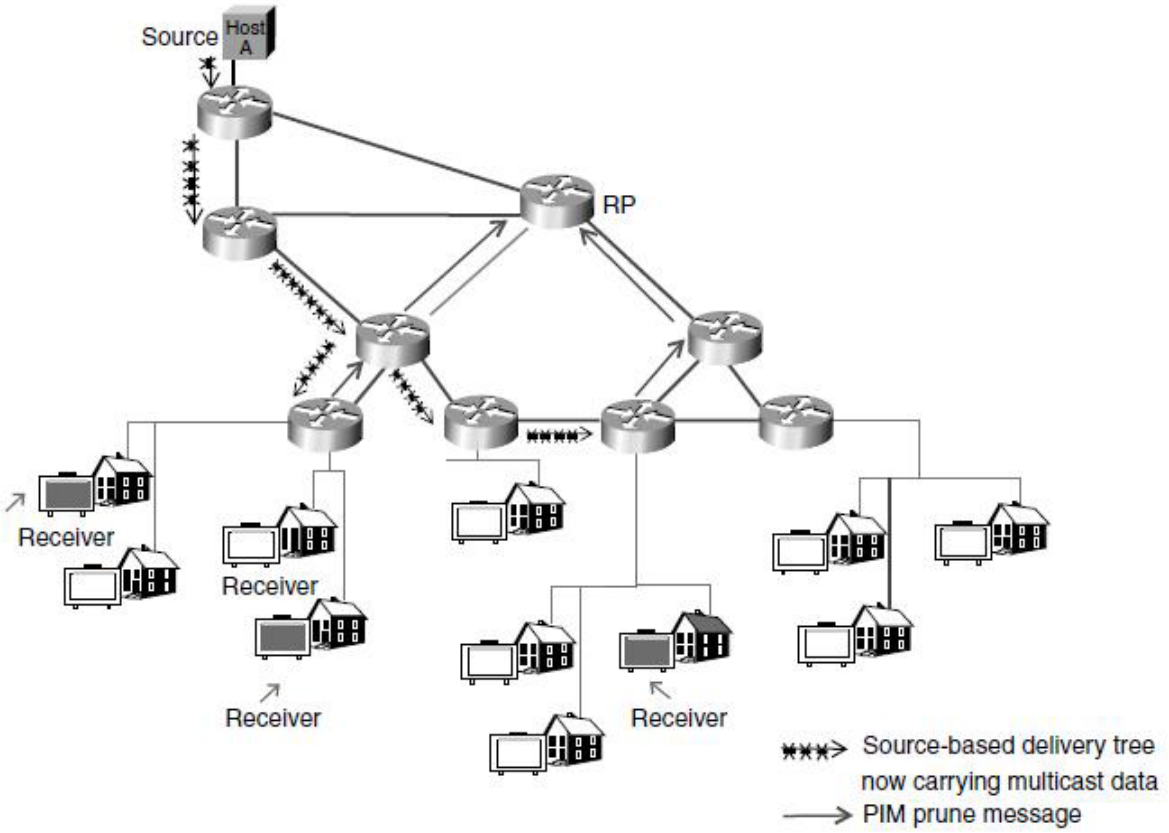
Harmadik fázis: SPT kiépítése a vevők felől

Elképzelhető, hogy az az útvonal, amit a csomagok a forrástól az RP-n át a vevőkig megtesznek, nem a legrövidebb. A csomagok fölösleges kerülő utat tesznek meg, ami miatt megnövekedik a késleltetés, és ez bizonyos alkalmazások esetében nem kívánatos lehet.

Ennek kiküszöbölésére a vevő DR-e kezdeményezheti egy forrás-specifikus SPT kiépítését a forrás irányába, kikerülvén ezzel az RP-t. Ehhez küld egy PIM (S, G) Join üzenetet az S forrásnak. Ez az üzenet eléri a forrás alhálózatát, vagy egy olyan routert, aminek van (S, G) bejegyzése. Ezután megindul az adatforgalom a forrástól a vevő fele, az SPT-n keresztül. A vevő ekkor még kétszer kap meg minden egyes csomagot, egyszer az SPT-n, és egyszer az RPT-n keresztül. Ezért a vevő DR-e küld egy (S, G) Prune üzenetet az RP felé. Ez az üzenet is végighaladva a közbenső routereken visszametszi a fának ezen ágát, így forgalom azon már nem fog érkezni a vevőhöz. [4]



11. ábra – Az SPT kiépülése a vevők felől



12. ábra – A fölös ágak visszametszése

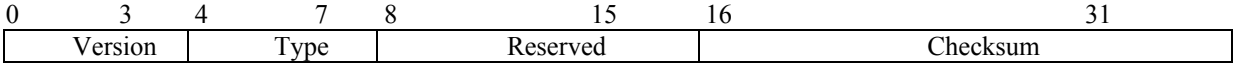
RP információk vétele

A Talákozási Pont (RP) IP címe egy domain összes routerében statikusan is beállítható, és dinamikusan is meghatározható. Az utóbbi esetben az RP információk szétküldéséhez úgynevezett Bootstrap üzeneteket használnak. A Bootstrap üzenetek végighaladnak a routereken keresztül a domainen belül, és a routerek ezekből szerzik az információt. Bootstrap üzenetek küldésére a Bootstrap router (BSR) hivatott, és az üzeneteket a 224.0.0.13 (ALL-PIM-ROUTERS) címre küldi. Ezek az üzenetek nemcsak az RP információk küldésére valók, hanem ezek segítségével választják meg a BSR-t is. A BSR megválasztása úgy történik, hogy a routerek egy kisebb csoportját BSR jelölt routernek (Candidate BSRs, továbbiakban C-BSRs) konfigurálnak, és egy egyszerű választó mechanizmussal megválasztják a BSR-t (a legnagyobb prioritás a nyerő). Szintén a routerek egy csoportját RP jelölt routernek (Candidate RPs, továbbiakban C-RPs) konfigurálnak. Ezek tipikusan ugyanazok a routerek, amik C-BSR-nek lettek beállítva. A C-RP-k periódikusan küldenek egy *C-RP-Értesítő*t (C-RP-Advertisement) a BSR-nek. Ezek a C-RP-Értesítők tartalmazzák magának a C-RP-nek a címét, valamint egy opcionális csoport címet, és a hozzá tartozó alhálózati maszkot. A BSR ezután beleteszi a Bootstrap üzenetekbe ezeket a C-RP-ket, a megfelelő csoport prefixumokkal.

A routerek veszik, és eltárolják ezeket a Bootstrap üzeneteket. Ha a DR kap egy Join üzenetet egy olyan csoporthoz, amihez még nincs bejegyzése, akkor elindít egy tördelő függvényt (hash function), amivel feltérképezi azt a C-RP-t, amelyiknek a csoport prefixébe beletartozik a kívánt csoport. Ha megvan, akkor továbbküldi a Join üzenetet annak az RP-nek.[18][1]

PIM üzenet felépítése

A PIM protokoll üzeneteinek közös fejrésze a következő:



13. ábra – PIM üzenetek közös fejrésze

A Típus (Type) mező adja meg, hogy pontosan mi a PIM üzenet célja, így csak ezt a mezőt fogjuk részletesen bemutatni, a fejrész többi része evidens.

Típus mező:

Típus	Leírás
0	Hello
1	Register
2	Register-Stop
3	Join/Prune
4	Boostrap
5	Assert
6	Graft (csak PIM-DM esetén használatos)
7	Graft-ACK (csak PIM-DM esetén használatos)
8	Candidate-RP-advertisement

Az egyes üzenetek felépítésének pontos bemutatása meghaladná ennek az összefoglalónak a kereteit (érdeklődők megtalálják az RFC 2362 4. részében: "Packet formats" címmel). Megjegyezzük, hogy amint a fenti táblázatban is látható, a Join és Prune üzenetek egyetlen közös típuskódot használnak, és az üzenet törzsében vannak felsorolva a mind a csatlakozni (Join), mind az elhagyni (Prune) kívánt források.

Hello. Ezen üzenetekkel térképezik fel a routerek, hogy melyik interfészükön vannak szomszédos PIM routerek, és választanak DR-t. Ezeket a 224.0.0.13 multicast címre küldik (ALL-PIM-ROUTERS). Ha a router kap egy ilyen üzenetet, akkor eltárolja a küldő IP címét, és dönt a DR felől. Mindig a legnagyobb IP című router lesz a DR. A Hello üzenet tartalmaz egy úgynevezett „visszatartási időt" (Hello-Holdtime), amennyi ideig a vevőnek érvényesként kell kezelnie az üzenetet. Ha a DR kiesik a rendszerből (a visszatartási idő lejár, és nem érkezik felőle újabb Hello üzenet), akkor egész egyszerűen új DR-t választ az adott router, úgy, hogy végignézi az interfészeit, és a legnagyobb IP című PIM router lesz az új DR.

Register/Register-Stop. Fentebb már említve volt a Register/Register-Stop üzenetek feladata. Amikor egy forrás először kezd el küldeni, akkor a forrás DR-e Register üzenetbe ágyazza a multicast csomagokat, és unicast módon küldi el az RP-nek. Az RP veszi ezeket az üzeneteket, és kiépíti a forrás-specifikus SPT-t a forrás felé. Miután kiépült az SPT, az RP küld egy Register-Stop üzenetet a DR felé, hogy többé már nincs szüksége Register üzenetekre.

Join/Prune. A Join/Prune üzenetek az elosztási fákhoz való csatlakozást, vagy kilépést teszik lehetővé a routerek számára. Ha egy DR kap egy csoporthoz való csatlakozási kérelmet (azaz IGMP Report üzentet) egy hozzá kapcsolódó hosttól, akkor küld egy PIM Join üzenetet az RP felé. Az összes router, amin az üzenet áthalad a DR és az RP közötti útvonalon, bejegyzi a (S, G) párost, és kiépül az osztási fa. Ha pedig a vevő nem kívánja tovább venni az adatforgalmat, akkor küld egy IGMP Leave üzenetet. A DR megkapja ezt a Leave üzenet, és küld egy PIM Prune üzenetet a PR felé, és a fa visszametsződik addig a pontig, ahol már vannak csoporttagok az ágak végén.

Bootstrap. A Bootstrap üzenet funkcióját már ismertettük az „RP információk vétele" részben.

Assert. Akkor használatos, ha egy adott ponthoz több útvonal is lehetséges. Az Assert üzenet tartalmaz egy távolság értéket (metric), ami az adott router távolságát jelzi az adott ponttól. Ha egy router a kimeneti interfészén kap csomagot, akkor küld egy Assert üzenetet, szintén a 224.0.0.13 címre, az ő távolság értékével. Ha az ő távolság értéke a legkisebb (azaz ő van legközelebb a forráshoz), akkor ő lesz a továbbító (forwarder), aki felé a downstream routerek küldeni fogják a maguk PIM üzeneteiket. Egyező távolság értékek esetén, a nagyobb IP cím a nyerő.

Candidate-RP-Advertisement. Bizonyos routereket RP jelöltként konfigurálnak. A tényleges RP ezen routerek közül kerül ki. A C-RP-k küldenek egy C-RP-Advertisement üzenetet a BSR-nek, amiben benne van a saját IP címük, és egy csoportcím lista, a prefix-el együtt. Ezzel határozza meg, hogy mekkora tartománynak lesz az RP-je. A BSR ezt beteszi a Bootstrap üzenetbe, és szétküldi a hálózaton, a routerek pedig ebből tájékozódnak az RP-t illetően. [3]

3.2.2 Protocol Independent Multicast – Dense Mode (PIM-DM)

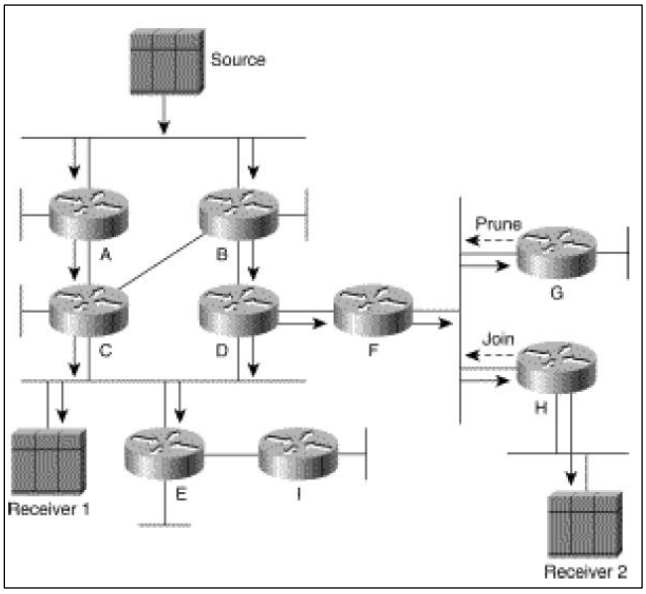
RFC 3973

A PIM-DM alapvetően akkor használatos, ha a hálózatban a multicast csoportok tagjai sűrűn helyezkednek el. Mivel ez nagyobb adatforgalmat is feltételez, mint a PIM-SM esetében, ezért nem jelent jelentősebb forgalomnövekedést, ha úgynevezett árasztásos módszerrel épül ki a fa.

Az árasztásos módszer lényege, hogy a beérkező csomagokat minden interfészen továbbküldjük (kivéve azon, amelyiken kaptuk). Ha a fa valamelyik ágán nincsenek vevők, akkor Prune üzenetekkel visszametsszük az ágat.

Ez a visszametszett állapot csak véges időtartamú, az adatszórás egy idő után újraindul (Cisco IOS-ek 3 percet definiálnak [9]). Ez alacsonyabb sávszélességű multicast forgalomnál még megengedhető, de magasabbnál már nem. Ezért egy úgynevezett állapot frissítő (State Refresh) üzenetet használnak annak meggátolására, hogy az adatfolyam magától újrainduljon. Ezt 60 másodpercenként küldik periodikusan. [8][9] Az adatforgalom újraindítására azért van szükség, mert ha egy router kiesik a rendszerből, és nem küld frissítő üzenetet, akkor se álljon meg véglegesen a forgalom.

A Prune üzeneteknek van egy három másodperces késleltetése, hogy adott esetben felül lehessen írni. Ahogy a 14. ábrán is látszik, a G routernek nincsenek vevői, míg az ugyanazon a LAN szegmensen lévő H routernek igen. Ezért a G router küld egy Prune üzenetet (224.0.0.13, All-PIM-Router címre), míg a H router ezt fogadván egy Join üzenetet küld. Így a H router meggátolta, hogy megszakadjon a multicast forgalom.



14. ábra – G Prune üzenete után H Join-t küld

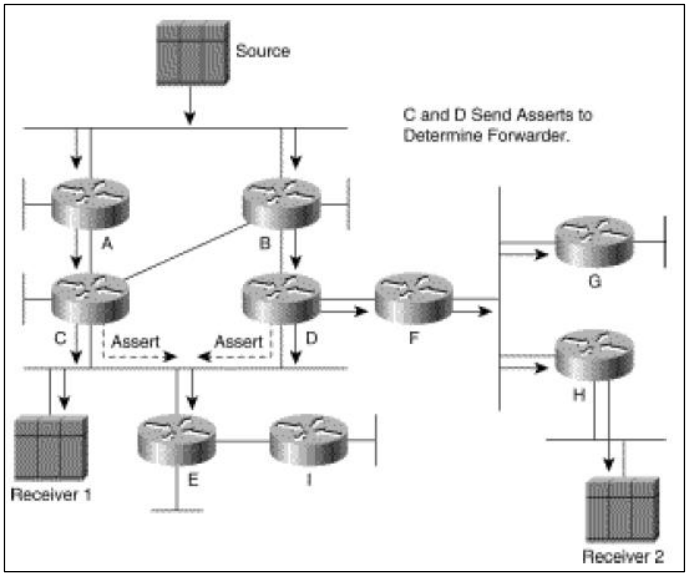
Prune üzeneteket egyébként az alábbi esetekben is küld a PIM router:

- Ha a forgalom nem az RPF (Reverse Path Forwarding) interfészen érkezik.
- Ha a faágak végén lévő routerhez nem kapcsolódnak közvetlenül vevők.
- Egy a faágon lévő router kap egy Prune üzenetet egy szomszédos routertől.
- Egy LAN szegmensen lévő router (akihez nem csatlakoznak közvetlenül vevők), kap egy Prune üzenetet egy ugyanazon LAN szegmensen lévő routertől, és nem kap fölülíró Join üzenetet.

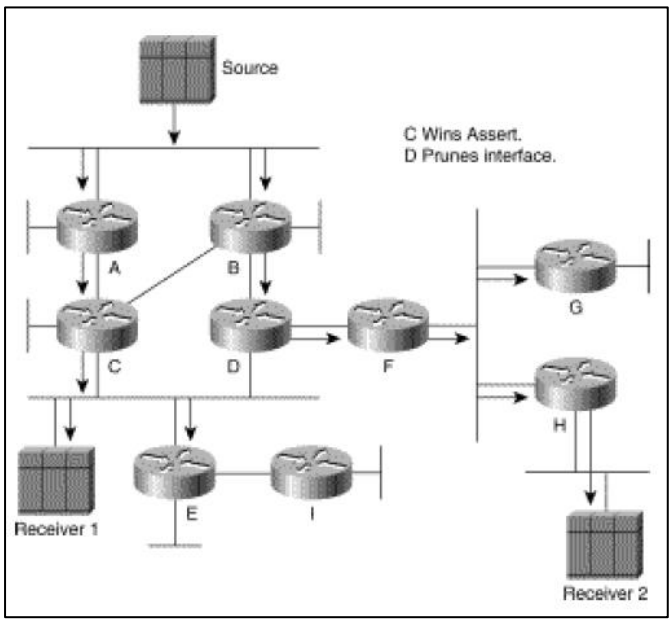
A PIM-DM is Assert üzeneteket használ a fölös, párhuzamos útvonalak kiküszöbölésére. Az Assert üzenet tartalmazza a router forrástól való távolság értékét, így eldönthető a többi

router számára, hogy ki van közelebb a forráshoz, és ki továbbítson (15. és 16. ábra szemlélteti).

Ha egy visszametszett ágon újra megjelennek a vevők, akkor a Graft üzenet segítségével újraaktiválhatjuk a korábban lemetszett ágat. Ezzel újraindíthatjuk az adatfolyamot, és nem kell megvárni, amíg letelik a három perc. Az a router, aki megkapta a Graft üzenetet, egy Graft-Ack igazoló üzenetet küld vissza így megbízhatóbbá téve a mechanizmust.



15. ábra – C és D megadja, milyen messze van a forrástól



16. ábra – C van közelebb, *ő* fog továbbítani, D nem

Irodalomjegyzék

[1] Beau Williamson - Cisco Press Publications – Developing IP Multicast Networks

[2] Lencse Gábor (2008) – Számítógép-hálózatok. UNIVERSITAS-Győr Nonprofit Kft., 2. kiadás.

[3] Daniel Minoli (2008) – IP MULTICAST with APPLICATIONS to IPTV and MOBILE DVB-H, John Wiley & Sons, INC., Publication, Canada.

[4] Xorp Inc. (2009, Január 7.) - Xorp User Manual Version 1.6 http://www.xorp.org/releases/current/docs/user_manual/user_manual.pdf Letöltve: 2010. 05. 23.

[5] Tom Smith - PIM Register Message Processing (2009. December 4.) http://technology-document.blogspot.com/2009/12/45-pim-register-message-processing.html Letöltve: 2010. 06. 10.

[6] Jákó András - Multicast routing napjainkban http://tiszai.tricon.hu/Tutor/lect01/IPMulticastRoutingNapjainkban.pdf Letöltve: 2010. 06. 15.

[7] Aurelian Pop - Marius Vlad (2002. 09. 30.) – Distance Vector Routing Protocol http://www.cs.tut.fi/~83390/syksy02/materials/DVMRP.pdf Letöltve: 2010. 07. 21.

[8] Peter J. Welcher (2001. October 01.) – PIM Dense Mode http://www.netcraftsmen.net/resources/archived-articles/376-pim-dense-mode.html Letöltve: 2010. 07. 29.

[9] http://www.cisco.com/en/US/docs/ios/12_1t/12_1t5/feature/guide/dtdmstrf.html Letöltve: 2010. 07. 29.

[10] Turányi Zoltán Richárd (1996) – Hálózati trendek http://www.szabilinux.hu/trendek/trendek74.html Letöltve: 2010. 06. 11.

[11] RFC 2236 (1997) – Internet Group Management Protocol, Version 2 http://www.faqs.org/rfcs/rfc2236.html Letöltve: 2010. 05. 02.

[12] RFC 1112 (1989) – Internet Group Management Protocol, Version 1 http://www.faqs.org/rfcs/rfc1112.html Letöltve: 2010. 05. 02.

[13] RFC 3376 (2002) - Internet Group Management Protocol, Version 3 http://www.faqs.org/rfcs/rfc3376.html Letöltve: 2010. 05. 02.

[14] IGMP Snooping www.h3c.com/portal/download.do?id=109048 Letöltve: 2010. 05. 22.

[15] Introduction to IGMP for IPTV networks (2007. October)
http://www.juniper.net/solutions/literature/white_papers/200188.pdf
Letöltve: 2010. 05. 02.

[16] http://www.firewall.cx/multicast-intro.php Letöltve: 2010. 04. 22.

[17] http://en.wikipedia.org/wiki/Real-time_Transport_Protocol Letöltve: 2010. 10. 02.

[18] Javvin Technologies, Inc.(2004-2005) – Network Protocols Handbook, 2. kiadás
http://books.google.com/books?id=D_GrQa2ZcLwC&pg=PA144#v=onepage&q&f=false Letöltve: 2010. 10. 4.

[19] RFC 3551 (2003. Június) – RTP Profile for Audio and Video Conferences with Minimal Control
http://tools.ietf.org/html/rfc3551 Letöltve: 2010. 09. 05.

[20] RFC 3550 (2003. Június) – RTP: A Transport Protocol for Real-Time Applications
http://tools.ietf.org/html/rfc3550 Letöltve: 2010. 09. 05.

[21] Larry L. Peterson, Bruce S. Davie (2007) – Computer Networks: A system approach, 4. kiadás
http://books.google.hu/books?id=zGVVuO-6w3IC&pg=PA430&dq=computer+networks+RTP&hl=hu&ei=we-6TLLbAoiH4Qbew6DNDg&sa=X&oi=book_result&ct=result&resnum=3&ved=0CDoQ6AEwAg#v=onepage&q=computer%20networks%20RTP&f=false
Letöltve: 2010. 09. 17.

[22] http://en.wikipedia.org/wiki/RTP_Control_Protocol Letöltve: 2010. 10. 06.

[23] RFC 2974 (2000. Október) – Session Announcement Protocol
http://www.faqs.org/rfcs/rfc2974.html Letöltve: 2010. 10. 02.

[24] RFC 4566 (2006. Július) – Session Description Protocol
http://tools.ietf.org/html/rfc4566 Letöltve: 2010. 10. 02.

[25] http://en.wikipedia.org/wiki/Session_Description_Protocol Letöltve: 2010. 10. 10.

[26] http://en.wikipedia.org/wiki/Real-time_Streaming_Protocol Letölve: 2010. 10. 11.

[27] http://www.cl.cam.ac.uk/~jac22/books/mm/book/node314.html Letöltve: 2010. 10. 11.

[28] RFC 2326 (1998. Április) – Real-time Streaming Protocol
http://www.ietf.org/rfc/rfc2326.txt Letöltve: 2010. 10. 11.

[29] Károly Dávid (2008) – IPTV és VoD rendszerek fejlesztése, szakdolgozat, Eötvös Loránd Tudományegyetem, Programozás elmélet és Szoftvertechnológia Tanszék http://karolyd.web.elte.hu/diploma/thesis.pdf Letöltve: 2010. 10. 01.

[30] IPv4 Multicast Address Space Registry (2010. 09. 22.) http://www.iana.org/assignments/multicast-addresses/multicast-addresses.xml Letöltve: 2010. 10. 12.

[31] David Malone, Niall Murphy (2005. Március) – IPv6 Network Administration, O'Reilly Publisher

[32] Implementing IPv6 Multicast http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-multicast.html#wp1055643 Letöltve: 2010. 10. 13.

[33] RFC 2710 – Multicast Listener Discovery (MLD) (1999. Október) http://www.faqs.org/rfcs/rfc2710.html Letöltve: 2010. 10. 13.

[34] RFC 1584 (1994. Március) – Multicast Extensions to OSPF http://www.faqs.org/rfcs/rfc1584.html Letöltve: 2010. 09. 17.

[35] Dudics Zsolt(2008) – Aktuális hálózati problémák megoldásainak vizsgálata – IP Multicast, Szakdolgozat, Debreceni Egyetem, Informatikai Rendszerek és Hálózatok Tanszék http://ganymedes.lib.unideb.hu:8080/dea/bitstream/2437/52004/1/Szakdolgozat.pdf Letöltve: 2010. 06. 19.

[36] IPv6 Multicast Address Space Registry (2010. 09. 13.) http://www.iana.org/assignments/ipv6-multicast-addresses/ipv6-multicast-addresses.xhtml Letöltve: 2010. 09. 13.

[37] Multicast in a Campus Network: CGMP and IGMP Snooping (2008. December. 17.) http://www.cisco.com/en/US/products/hw/switches/ps708/products_tech_note09186a00800b0871.shtml#igmp_snooping Letöltve: 2010. 09.17.