

IPTV protokollok

oktatási segédanyag az „IP alapú távközlés” c. tárgyhoz

Készült:

Steierlein Balázs:

IPTV rendszerek

című szakdolgozatának felhasználásával.

Szerkesztette: Lencse Gábor

FIGYELEM! Ez a segédanyag a szakdolgozat félkész munkaverziója alapján készült, egyes részei pontatlanok, befejezetlenek lehetnek. A segédanyag a tárgyalt protokollok céljának és működési elvének bemutatására szolgál; a protokollok minden részletének pontos megismerésére nem feltétlenül alkalmas!

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

Az anyag témakörei:

IP (v4, v6) multicast fogalma, IGMP, IGMP snooping, MLD.

Multicast routing protokollok közül: PIM-SM, PIM-DM

Tartalom

1. IP MULTICAST	3
1.1 MULTICAST HÁLÓZATI CÍMZÉS IPv4-ES ÉS IPv6-OS KÖRNYEZETBEN	4
1.2 MULTICAST ETHERNET CÍMZÉS IPv4-ES ÉS IPv6-OS KÖRNYEZETBEN	5
2. CSOPORTMENEDZSMENT	6
2.1 INTERNET GROUP MANAGEMENT PROTOCOL (IGMP)	6
2.1.1 IGMP üzenet felépítése.....	7
2.1.2 Állapot átmeneti diagramok	8
2.1.3 IGMP Snooping	9
2.2 MULTICAST LISTENER DISCOVERY (MLD).....	11
3. MULTICAST ROUTING PROTOKOLLOK.....	13
3.1.1 Protocol Independent Multicast – Sparse Mode (PIM-SM)	13
3.2.2 Protocol Independent Multicast – Dense Mode (PIM-DM)	18
IRODALOMJEGYZÉK	21

1. IP Multicast

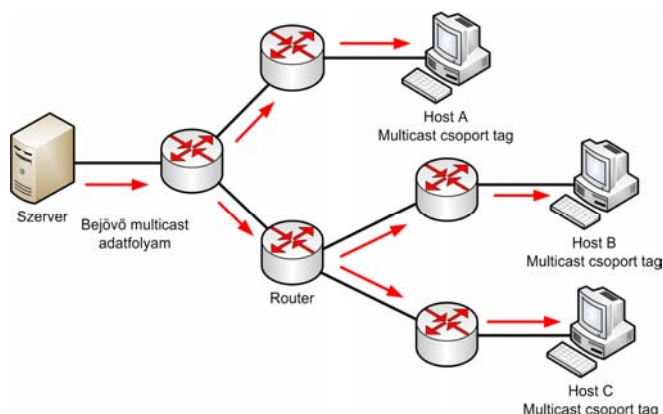
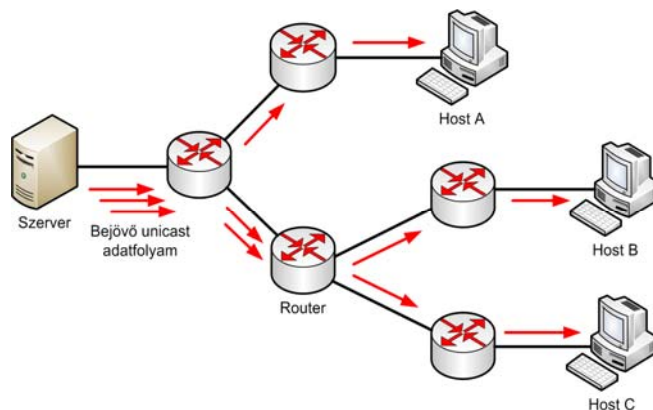
RFC 1112

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

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

Unicast címzés esetén pedig 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.

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 csoport címre (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.



1.1 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 tranzien csoport.

Az **állandó csoport** akkor sem szűnik meg, ha az utolsó tag is kilépett. Ezek a csoportok általában a routing protolloknak, és egyéb topológia felderítő mechanizmusoknak van fenntartva. A számunkra fontosabbakat az alábbi listában soroltam fel [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 **tranzien csoportok** rendszerint valamely feladat idejére jönnek létre, és ha kilépett az utolsó tag is, akkor megszűnnek.

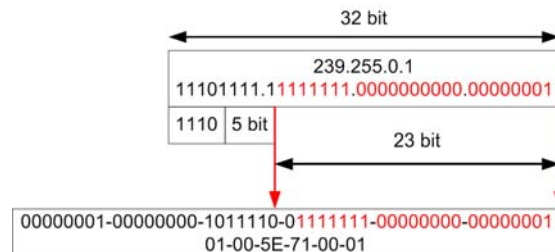
A 224.0.0.0/24 címtartomány úgynevezett link-local tartomány, az ide tartozó címekre küldött üzeneteket a routerek nem továbbítják. A privát tartományt bárki felhasználhatja a saját rendszerén belül (a unicast privát IP címekhez hasonlóan.)

IPv6-os környezetben az FF00::/8 prefixumtól 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

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.



A leképzés megértéséhez vizsgáljuk meg a MAC címek felépítését! Először is az OUI (Organizationally Unique Identifier) első¹ 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 3 bájtja előre meg van határozva: 01-00-5E. A második három bájtba pedig az IP címet illesztjük be. IPv4 esetén első ránézésre problémát jelenthet, hogy az IP cím 4 bájtos (vagyis 32 bites). Ezt úgy oldják meg, hogy mivel minden multicast IP cím az 1110 prefix-el kezdődik, így ezt elhagyhatjuk. Ez még mindig kevés, azonban történeti okok miatt még további 5 bitet el kell hagynunk, és így jutunk el a maradék 23 bitig, amit már berakhatunk a MAC címbe. Ennek hátránya, hogy mivel nem a teljes IP cím kerül leképezésre, így több multicast IP címnek lesz ugyanaz a MAC címe.[16]

IPv6-nál valamivel egyszerűbb a helyzet. A MAC cím első két bájtja fixen 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 elől) 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 1112, RFC 2236, RFC 3376

Az IGMP protokollt a hostok multicast csoporthoz való csatlakozásukhoz, ill. 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 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 egy all-hosts 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 group 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 lejártá után. A hostok a groupnak címezik a report üzenetet, így mindenki megkapja a csoportban. Így ha a host timere lejártá 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.

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 elkerülése érdekében, hogy ez a report nehogy elveszen, vagy megsérüljön, a hostnak javallott rövid időn belül megismételnie ezt az üzenetet.

Ha több IGMP képes router is van egy szegmensben, akkor egy úgynevezett „*IGMP Querier Election*” mechanizmust használnak annak eldöntésére, hogy ki legyen az, aki majd a querykat fogja küldeni (Querier vagy Non-Querier) a hálózaton. Ez egész egyszerűen úgy működik, hogy mindig a legalacsonyabb IP című router lesz a Querier. Ennek kiderítésére a routerek periódikusan küldenek általános queryket minden hozzájuk kapcsolódó hálózatra.

Minden IGMP képes router alaphól querierként kezd, és csak utána lépnek vissza non-querier állapotba.

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 a router kap egy ilyen Leave üzenetet, küld egy csoport-specifikus 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*) lejártá 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.

2.1.1 IGMP üzenet felépítése

0	7	8	15	16	31
Type		Max Resp Time			Checksum
Group Address					

1. IGMPv2 üzenet

Type: az üzenet típusa. Háromféle üzenet típus van: Membership Query, Membership Report, Leave Group. A Membership Querynek még további 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 üzenet típusnál az értéke nullára van beállítva, és a címzettek figyelmen kívül hagyják.

Checksum: ellenőrző összeg.

Group Address: Membership Query üzeneteknél, ha az 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.

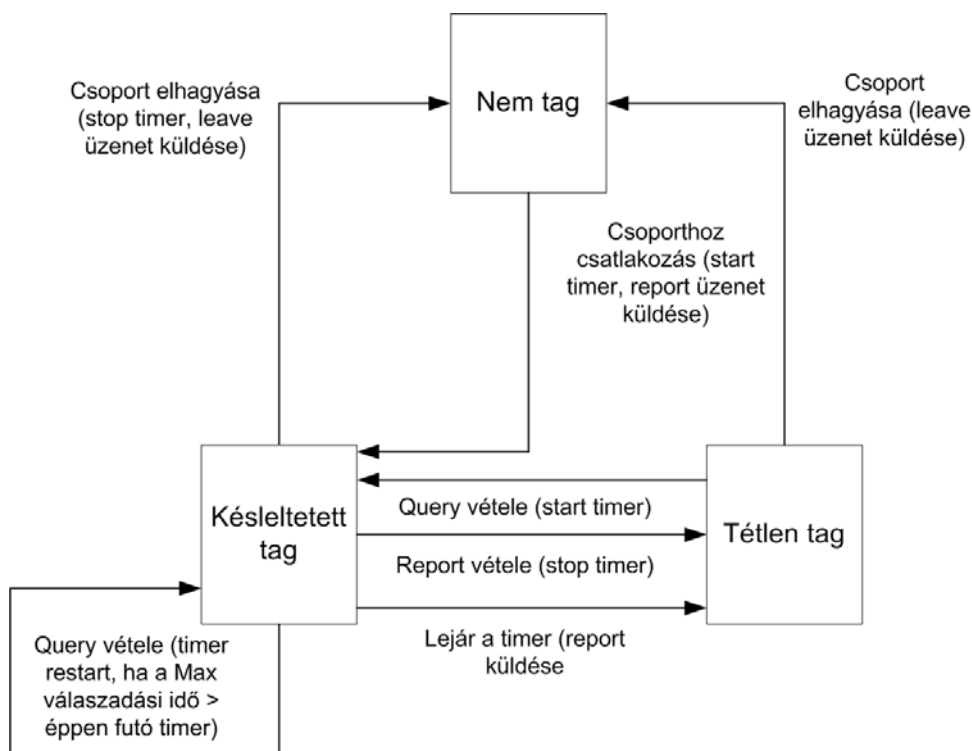
2.1.2 Állapot átmeneti diagramok

A host három állapotban lehet:

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

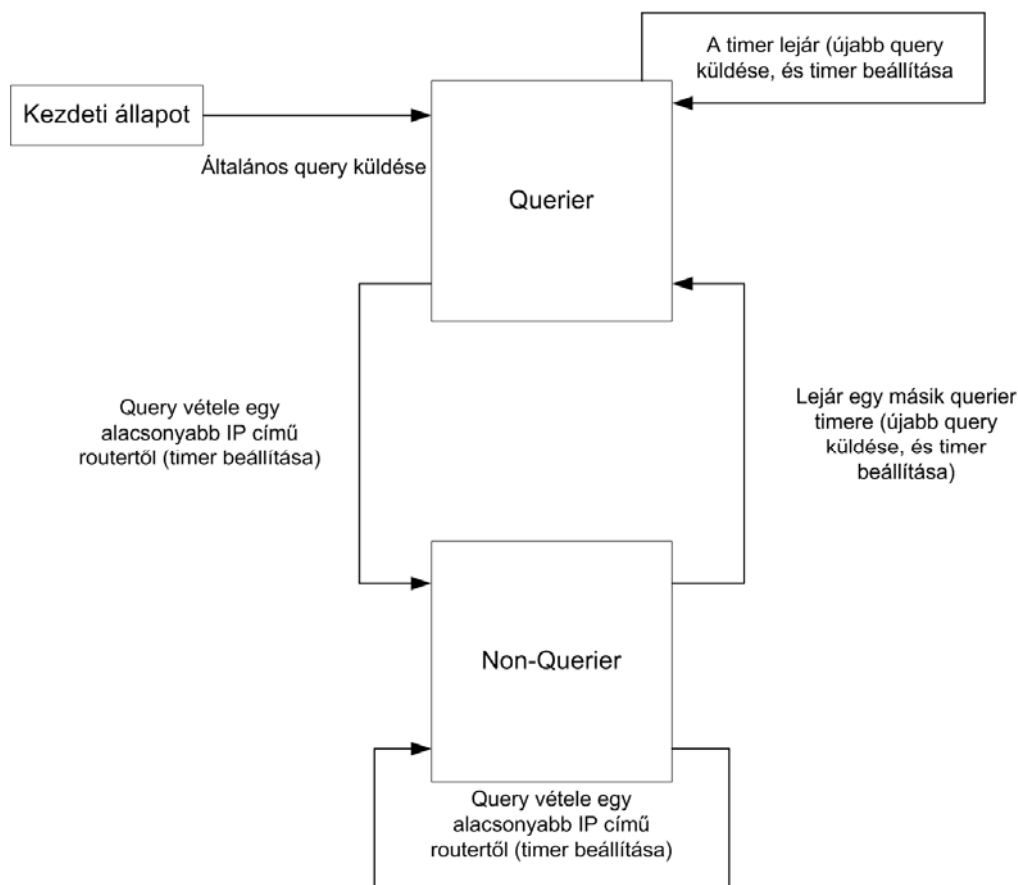
Egy router két állapotban lehet:

- Querier: a querier router van kijelölve az IGMP üzenetek küldésére.
- Non-Querier: ez(ek) a router(ek) nem küld(enek) query üzeneteket.



2. ábra A host állapotátmeneti diagramja

Az IGMPv2 és az IGMPv1 is az ASM (*Any Source Multicast*) 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. Az IPTV rendszerekben legáltalánosabban az IGMPv2-t használják, ami természetesen (az IGMPv3-al együtt) visszafele kompatibilis.

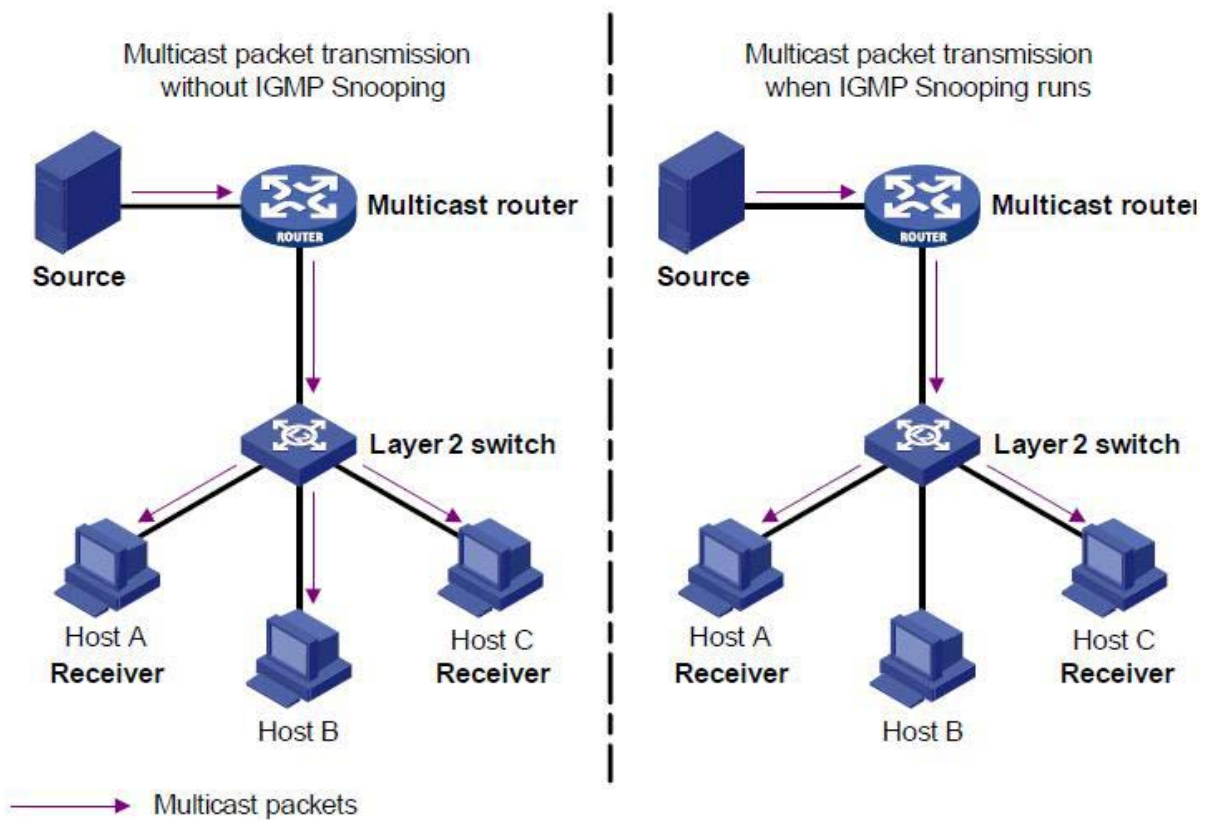


3. ábra A router állapotátmeneti diagramja

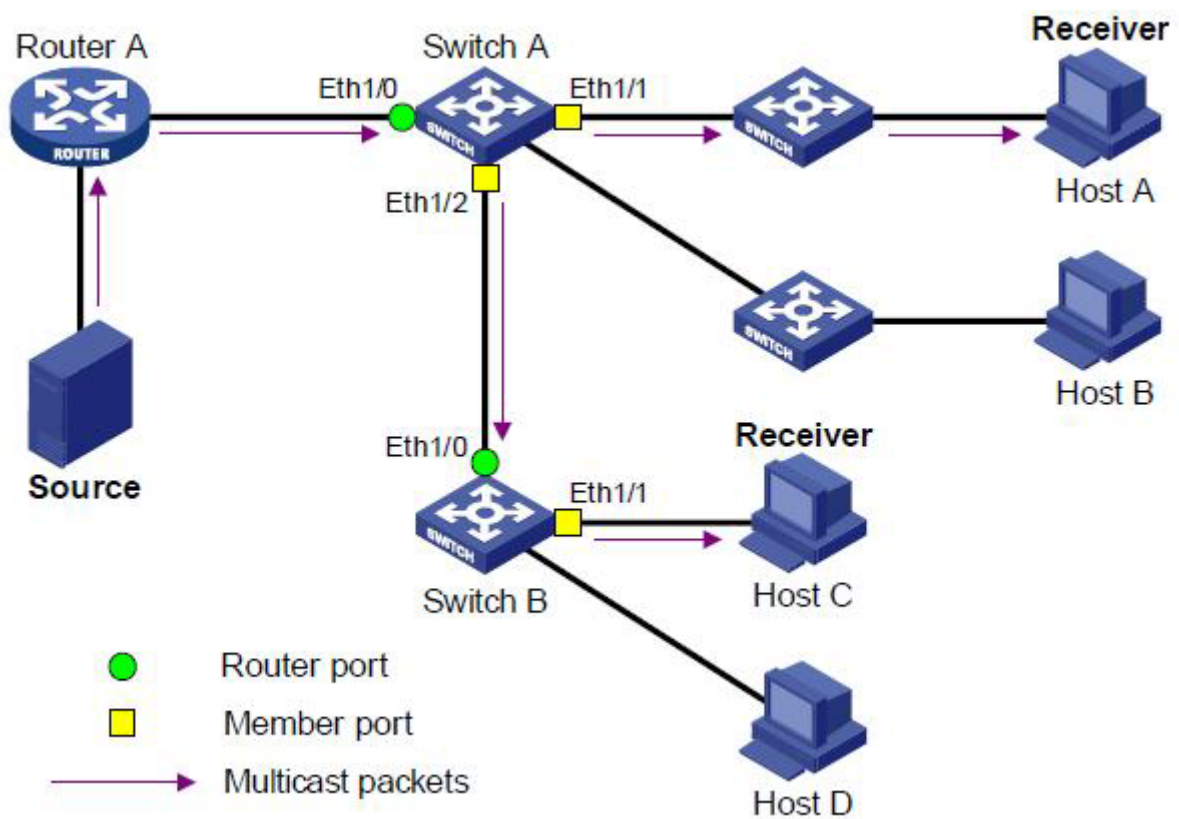
2.1.3 IGMP Snooping

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 protokollban mezőinek vizsgálatát, sőt a protokollban való részvételét teszi szükségessé!



4. IGMP Snooping



5. Router és tag portok

Működése

A könnyebb érthetőség és egyszerűség érdekében definiáljunk két fogalmat a switchekkel kapcsolatban: router port, és tag (member) port. A router port jelentse a switchnek azt a portját, ami a router irányába „mutat”, a tag port pedig azt, amin keresztülhaladva egy csoport tag felé közelednénk.

A switchnek mindenek előtt tisztában kell lennie azzal, hogy melyik portján vannak routerek, és melyeken kliensek, azaz csoporttagok.

A routerek megtalálásának érdekében következő üzeneteket figyel:

- 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)

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 szeretne. A csoporthoz pedig hozzárendeli az összes router portot, és azt 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 saját 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:

0	7 8	16 17	32
Type	Code	Checksum	
Maximum Response Delay		Reserved	
Multicast Address			

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 nullázza, a vevő figyelmen kívül hagyja.

Checksum. Ellenőrző összeg.

Maximum Response Delay. Maximum várakozási idő lényegében. Csak Query üzeneteknél értelmezzük, milliszekundumban. A többi üzenet fajtánál ez a mező nullázva van.

Reserved. Fenntartott mező, a forrás nullázza, a vevő figyelmen kívül hagyja.

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

3. Multicast Routing Protokollok

Azokban a hálózatokban ahol az adatok több routeren, vagy akár több alhálózaton keresztül jutnak el a forrástól a vevőig, routing protokollokra van szükség. 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 benne van, független multicast protokoll, ami azt jelenti, hogy nincs saját topológia felderítő algoritmus, hanem egy másik protokoll (pl. RIP, 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 fentebb 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*).

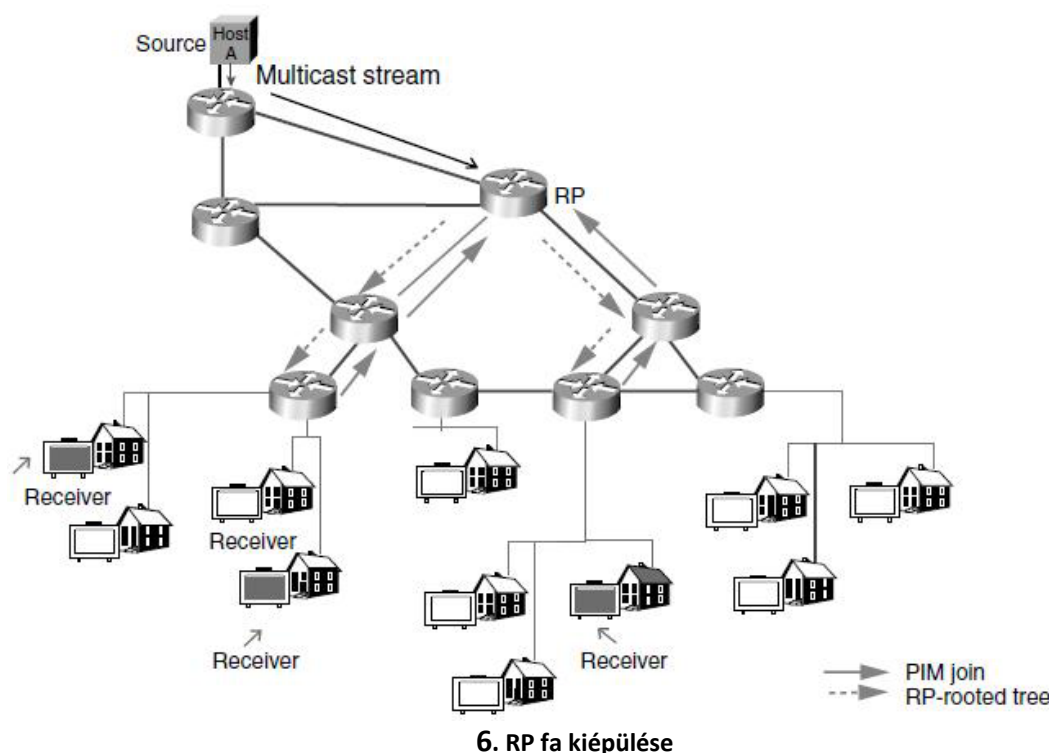
A PIM-SM protokoll működésének 3 fázisa van.

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. Az első fázisban a vevő jelzi igényét a multicast forgalomra, ami egy adott csoportnak van címezve. Ehhez IGMP, vagy MLD (Multicast Listener Discovery) protokollt használ, attól függően, hogy IPv4, vagy IPv6-os környezetről van szó. A vevő helyi PIM routerei közül választanak (lásd az apró betűvel írt részben: Hello üzenet leírása) egy *Kijelölt routert* (Designated Router, továbbiakban DR). Ha a DR megkapja a vevőtől a kérést, akkor küld egy

PIM Join üzenetet a kívánt multicast csoportnak a *Randevú Pont* felé (továbbiakban RP). Az RP egy PIM-SM router, ami a forrás és a vevők között segít kiépíteni a kapcsolatot. Az címét vagy statikusan beállítják az AS összes routerén, vagy egy megfelelő algoritmussal megválasztják az AS PIM routerei közül (lásd az apró betűvel írt részben: RP információk vétele).

Az előbb említett Join üzenetnek a jelölése: (*, G). Az első tag a forrás, a második tag 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. Ezt a fát nevezzük megosztott fának. Ha az összes 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.



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

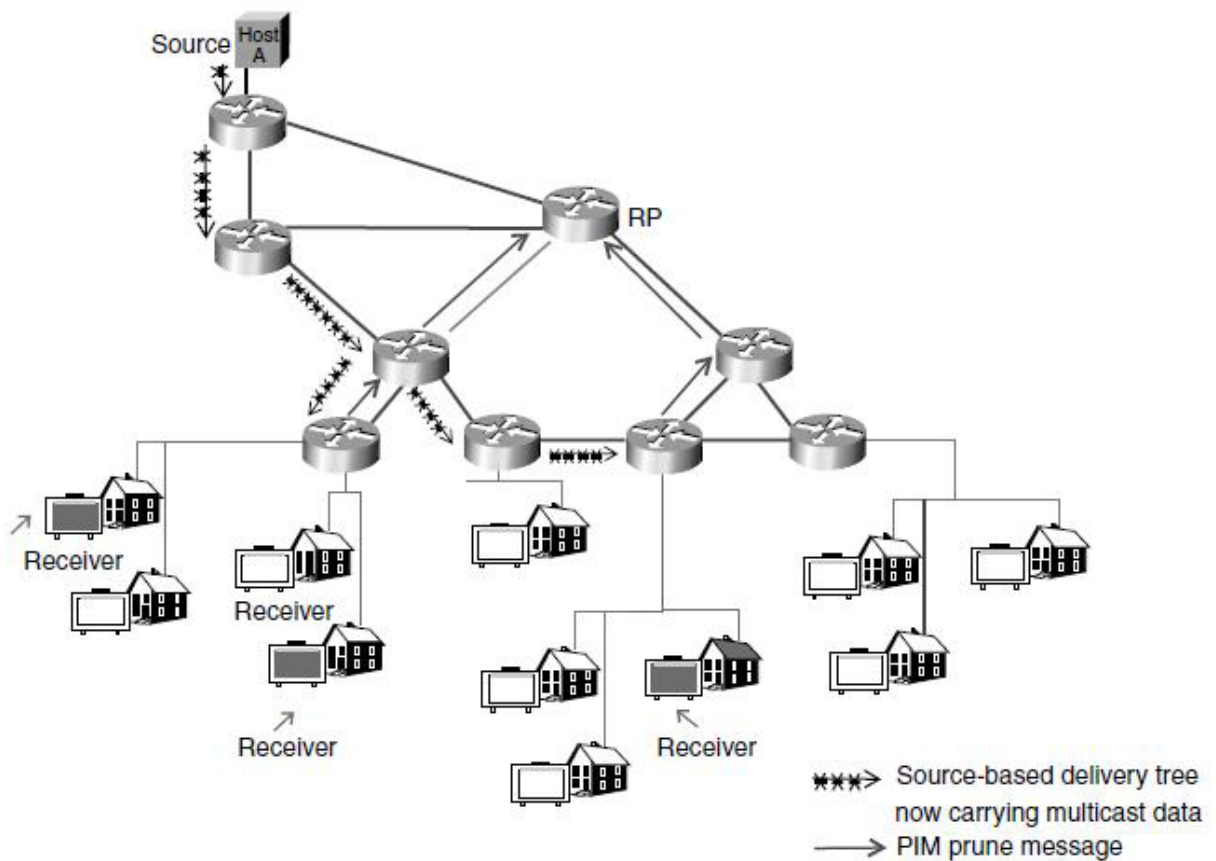
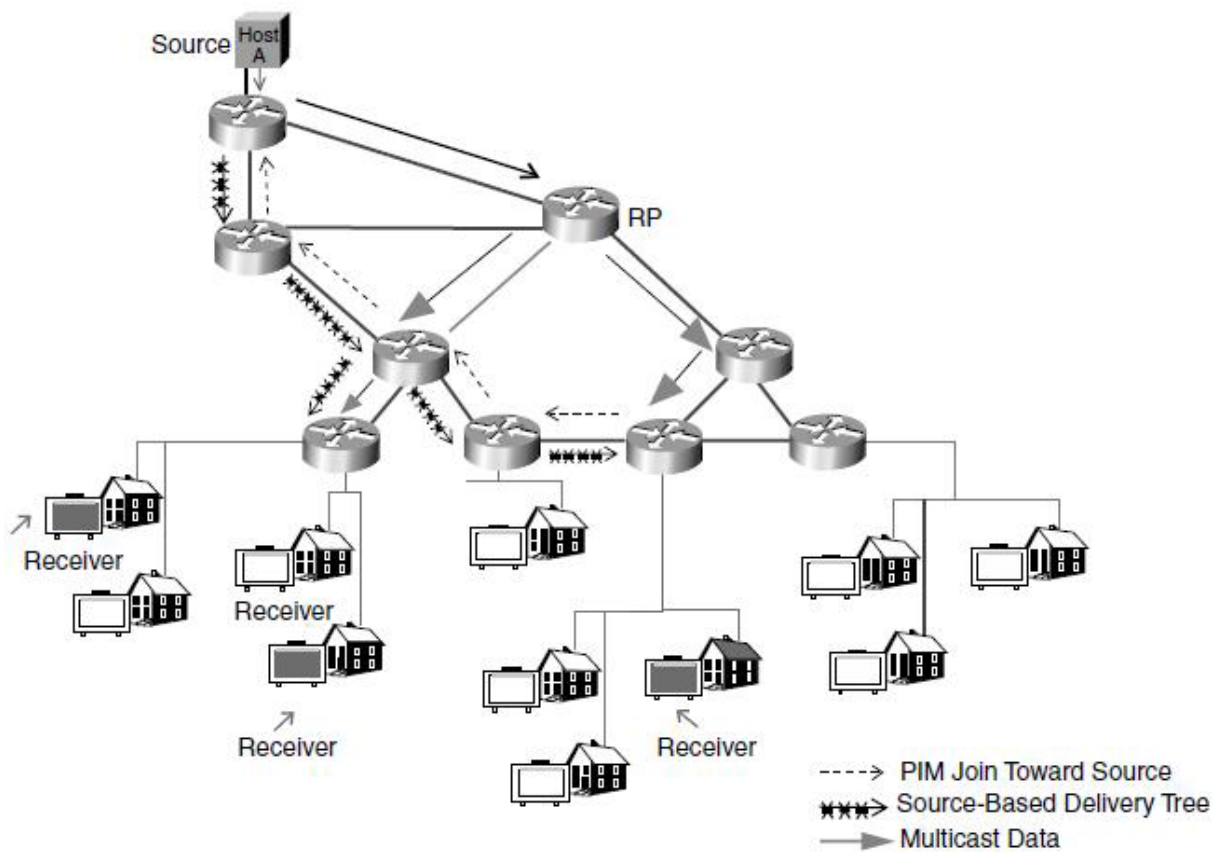
A források első multicast csomagjait a first-hop routerek unicast IP csomagokba ágyazzák és így küldik el az RP-nek. Ezeket a csomagokat hívjuk regiszter csomagoknak. A forrás ezzel az üzenettel tudatja az RP-vel, hogy készen áll adatfolyam küldésre. Az RP

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 (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 DR-ét vagy egy olyan routert, aminek már van (S,G) bejegyzése, akkor az adatok elkezdnek 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 vevőkhöz megtesznek, nem a legrövidebb. A csomagok fölösleges kerülő utat tesznek meg, ami miatt megnövekedik a várakozási idő, és ez bizonyos alkalmazások esetében nem kívánatos lehet.

Ennek kiküszöbölésére a vevő DR-e kezdeményezheti egy forrás-specifikus SPT kiépítését a forrás irányába, kikerülve ezzel az RP-t. Ehhez küld egy (S,G) Join üzenetet a forrásnak (S). Ez az üzenet eléri a forrás alhálózatát, vagy egy olyan routert, aminek van (S,G) bejegyzése. Ezután megindul az adatforgalom a forrástól a vevő felé, 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. Természetesen ez az üzenet is végighaladva a közbenső routereken visszametszi a fának ezen ágát, így forgalom azon már nem fog érkezni a vevőhöz. [4]



RP információk vétele

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 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 priorítás 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-Értesítőt (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-eket, 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

0	3	4	7	8	15	16	31
Version	Type	Reserved	Checksum	Data			

9. PIM üzenet felépítése

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

Típus mező:

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

Hello. Ezen üzenetekkel térképezik fel a routerek, hogy melyik interface-ükön 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 interfaceit, 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 router (pl. a DR) kap egy IGMP join üzenetet egy közvetlenül hozzá kapcsolódó host-tól, akkor küld egy PIM Join üzenetet az RP felé. Az összes router, amin áthalad a DR és az RP közötti útvonalon, bejegyzik 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 ismertettem az „RP információk vétele” részben, így arra még egyszer nem térek ki.

Assert. Akkor használatos, ha egy adott ponthoz több útvonal is lehetséges. Az Assert üzenet tartalmaz egy „metric” értéket, ami az adott router távolságát jelzi az adott ponttól. Ha egy router a kimeneti interfészen kap csomagot, akkor küld egy Assert üzenetet, szintén a 224.0.0.13 címre, az ő metric értékével. Ha az ő metric 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 metric 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

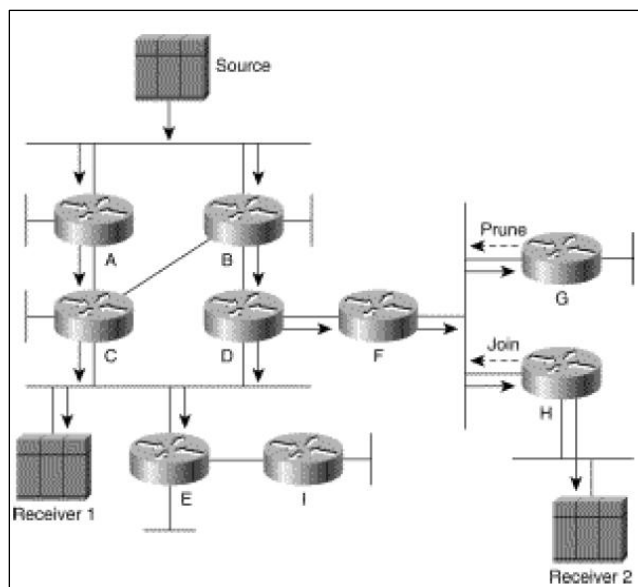
A PIM-DM alapvetően akkor használatos, ha a hálózatban a multicast csoportok tagjai sűrűn helyezkednek el. Mivel ez nagyobb adatforgalmat is feltételez, mint a PIM-SM esetében, ezért nem jelent jelentősebb forgalomnövekedést, ha úgynevezett árasztásos módszerrel épül ki a fa.

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 visszametszzük 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 10. 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.



10. ábra

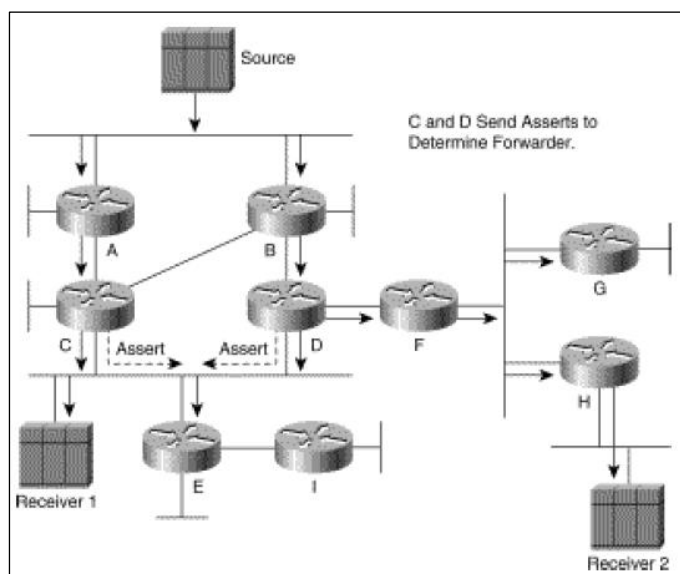
Prune üzeneteket egyébként az alábbi esetekben is küld a PIM router:

- 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 fölülrő Join üzenetet.

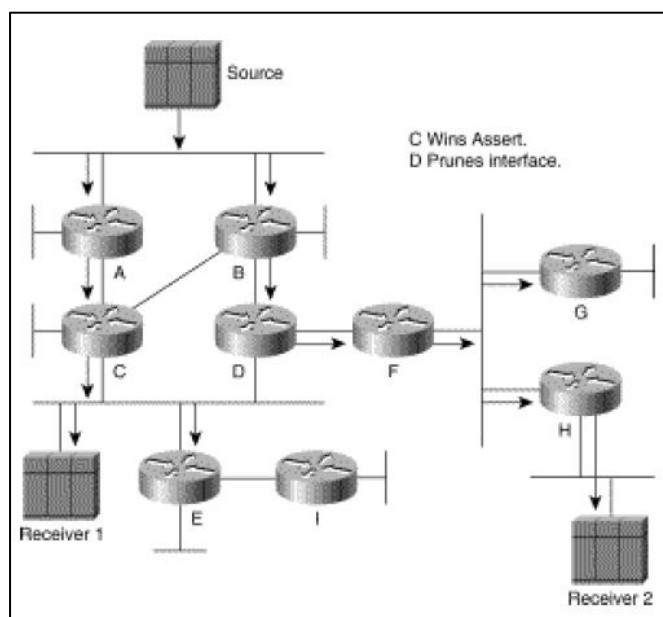
A PIM-DM is Assert üzeneteket használ a fölös, párhuzamos útvonalak kiküszöbölésére. Az Assert üzenet tartalmazza a router metric értékét (távolságát a forrástól), így eldönthető a többi router számára, hogy ki van közelebb a forráshoz, és ki továbbítson (11. és 12. ábra szemlélteti).

Ha egy visszametszett ágon újra megjelennek a vevők, akkor a Graft üzenet segítségével újraaktiválhatjuk a korábban lementszett 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.



11. ábra



12. ábra

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://tisza.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 Letöltve: 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 Szoftvertchnológia 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://ganyemedes.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.