



Budapesti Műszaki és Gazdaságtudományi Egyetem

Unicast és multicast szolgáltatást támogató IP maghálózatok tervezésének és analízisének támogatása

Diplomaterv –V4

6. fejezet Multicast Routing

Szerző: Mile Szabolcs Gyula

Konzulens: Jakab Tivadar

2007. november 02.

6 Multicast Routing

6.1 Bevezető

Háromféle csomagtovábbítási módszert különböztetünk meg, unicast, broadcast és multicast küldést.

Unicast (one-to-one) küldés esetén a csomagok egyetlen forrástól egyetlen nyelőhöz irányulnak. A teljes útvonal során az útvonalválasztást a routerek végzik router tábláik alapján.

Broadcast (one-to-many) esetben egy csomagot a forrástól a hálózat valamennyi entitásához elküldünk, a routerek intelligens módon továbbítják a bejövő csomagokat valamennyi kimenő irányba.

Multicast (one-to-many, many-to-many) küldés esetén a csomagokat csak a felhasználók egy specifikus csoportjának továbbítjuk.

Jelenleg még a legtöbbet használt csomagtovábbítási mód az unicast küldés, de a szolgáltatások fejlődésével egyre több olyan esettel találkozhatunk (Webinárok, Webrádió, IPTV, információs szolgáltatások, konferenciabeszélgetések stb...), mikor ugyanazt az információt több felhasználóhoz is el kell juttatni. Ekkor a hagyományos unicast módú továbbítás több ok miatt is kivitelezhetetlen lenne:

- A forrásnak nyilván kellene tartania valamennyi nyelőjét
- Ugyanazon IP csomag több példánya is terhelné a linkeket, mely hatalmas erőforrásigényhez vezetne

A fenti esetekre az optimális megoldást a multicast csomagtovábbítás jelenti. Jelentős előnyei mellett azonban meg kell említeni, hogy a multicast hálózati működés sokkal komplexebb, mint unicast illetve broadcast esetben. A küldés során számos hálózati entitás, illetve folyamat együttműködése szükséges. Sajnos a régi hálózati berendezések még nem támogatják a multicast üzemmódot, így ezek cseréje (pl. multicast router) is szükséges lehet, mely további költségeket ró a szolgáltatókra.

6.2 Multicast Csoportok

A fenti első problémára a megoldás, hogy az azonos igényű nyelőket egy csoportba rendezzük, s a csoportokhoz hozzárendelünk egy D osztályú IP címet. A forrásnak elég a különböző csoportokhoz tartozó IP címeket ismernie, és eltárolnia.

A multicast címek felhasználását is, mint az összes többi osztály esetében az IANA szervezet felügyeli. A D osztályú címek a 224.0.0.0 – 239.0.0.0 tartományig terjedő IP címeket foglalja magában, de ez a tartomány is több kisebb szegmensre van felosztva.

224.0.0.0 – 224.0.0.255 Routing protokollok számára

224.0.1.0 – 224.0.1.255 Internet controll block

224.0.2.0 – 238.255.255 Globális multicast címek

232.0.0.0 – 232.255.255.255 SSM címek lásd később!!!

239.0.0.0 – 239.255.255 Lokális multicast címek

A globális címek kiosztása központi regisztráción keresztül történik, míg a lokális címek egy AS-en belül szabad felhasználásúak. Lényegében a multicast kommunikáció első lépése a különböző szolgáltatások IP címhez kötése.

A layer 2-es rétegben szintén kijelöltek egy címtartományt, és a multicast címeket ezen címtartományra képezik le. A rendelkezésre álló multicast MAC tartomány (0100.5e00.0000 – 0100.537f.ffff) kisebb, mint a multicast IP tartomány, így egy MAC címnek valójában 32

multicast osztály felel meg. Az egyértelmű leképzés végül a felsőbb rétegek segítségével történik.

6.3 Any-source multicast vs. Source specific Multicast

A multicast kommunikáció alapgondolata Stephen Deeringtől származik. Ő a következőképpen definiálta a multicast működését:

IP szemantika: Egy forrás bármikor, bármilyen előfeltevés nélkül tud multicast adatokat küldeni, ehhez nem szükséges sehol regisztrálnia. Az adatcsomagokat UDP protokollon továbbítják.

Nyitott csoportok: A forrásoknak csak egy multicast címet kell ismerniük. Nem szükséges a csoportról információval rendelkezniük, s egyáltalán nem szükséges, hogy maguk is csoporttagok legyenek. Egy csoportnak bármennyi forrása lehet.

Dinamikus csoportok: A tagok bármikor szabadon csatlakozhatnak, és szabadon hagyhatják el a csoportot.

Ezt a fajta működési módot nevezzük any source multicastnak (ASM). Jelölése: (*,G)

A nyelő csak a csoport címével azonosítja a multicast streamet. Az idők folyamán azonban nyilvánvalóvá vált, hogy különösen biztonsági szempontból (DoS) szerencsésebb lenne egy olyan működési mód is, ahol a nyelők meg tudják adni, hogy kikkel szeretnének kommunikálni, és csak a tőlük származó streamet kapját meg. Ezt nevezzük source specific multicastnak (SSM). Jelölése: (S,G). A csoport mellett a forrás címének megadása is kötelező.

6.4 Multicast Group Membership Discovery Protokoll

Ha egy nyelő szeretne csatlakozni egy adott multicast csoporthoz, akkor regisztrálnia kell magát az adott hálózat multicast DR-nél. IPv4 esetén ez a regisztráció IGMP üzenetek segítségével történik, míg az IPv6 hálózatokban az MLD protokollt használják.

6.4.1 Internet Group Management

Ha egy nyelő csatlakozni szeretne egy csoporthoz, akkor egy IGMP regisztrációs üzenetet küld a LAN DR-ének. Amennyiben a router még nem csoporttag, akkor valamilyen multicast protokoll segítségével regisztrálja magát. Ha csoporttag volt, akkor folytatja eddigi működését. Egy speciális entitás, az IGMP Querier mely maga is egy MC Router periodikusan lekérdező üzeneteket küld a hálózaton. Ezen üzenetekre adott válasz alapján a multicast routerek számon tudják tartani, hogy milyen csoporttagok találhatóak a hálózaton. Ha van olyan csoport, melynek már nincs aktív tagja a LAN-on, akkor a routerek a használt multicast protokoll segítségével leiratkoznak a csoporttagságról.

Az regisztrációs üzenet lényegében egy hagyományos 1-es TTL-ű IP datagramm, melynek a célcím mezője tartalmazza a multicast csoport címét.

Az IGMPv1-ben a querier kiválasztása a multicast routing protokolltól függött, míg a második verzió már tartalmaz egy kiválasztási mechanizmust. Egy router alapértelmezettként egyben IGMP querier is. Ha kap egy üzenetet arról a hálózatról, ahol még querier, és a küldő router kisebb IP címmel rendelkezik, akkor nem lesz tovább querier. Ha egy meghatározott ideig nem kap query üzenetet, akkor megint querier státuszba lép.

Mivel a multicast routerek csak binárisan tartják nyilván a szegmensükön lévő csoportokat, vagyis, hogy egy adott csoport létezik vagy sem, így optimalizálni lehet a query üzenetekre adott válaszok számát. Mikor a nyelők megkapják a query üzenetet, akkor csoportonként egy véletlenszerű időzítőt indítanak. Ha az időzítő lejár, akkor elküldik a regisztráló üzenetet a routernek. Mivel valamennyi nyelő megkapja a regisztráló üzenetet, így detektálni tudják, hogy egy másik nyelő már visszajelzett-e egy adott csoportra. Ha egy

csoportra már történt visszajelzés, akkor nem kell több regisztráló üzenetet küldeni, így csökkenthető a küldendő üzenetek száma.

Az IGMPv2-ben bevezetésre került egy Leave Group üzenet, mely segítségével jelezni tudja egy csoport tagja, hogy nem szeretne többé csoportspecifikus üzenetet kapni. Ez azért előnyös, mert nem kell megvárni a query time végét egy adott stream megszüntetéséhez. Amennyiben ilyen üzenetet kap a router, akkor egy csoportspecifikus kérést indít a hálózaton. Ha nem érkezik rá regisztrációs válasz, akkor megszünteti a stream továbbítását a helyi hálózatra, és maga is leiratkozik a streamről.

Az IGMP protokoll harmadik verziójában bevezetésre került a forrásszűrés. Ez az eljárás lehetővé teszi, hogy a regisztrációs üzenettel küldhessünk egy engedélyező-, illetve tiltó listát, mely tartalmazza, hogy mely forrásoktól szeretnénk adatokat kapni, illetve hogy melyektől nem. Erre ez eljárásra a SSM támogatása miatt volt szükség.

6.4.2 Multicast Listener Discovery

Jelenleg két MLD verzió használatos. Az első verzió működése megegyezik az IGMPv2-vel, míg a 2. verzió teljesen azonos az IGMPv3-al.

6.5 Multicast Routing Protokoll

A multicast küldés elengedhetetlen feltétele, hogy az egyes natív multicast routingot támogató berendezések képesek legyenek egymással kommunikálni, vezérelni a különböző adatfolyamokat. Ezt biztosítják a kizárólag a routerek között használt multicast routing protokollok.

A téma komplexitása és fontossága miatt az egyes routing protokollok egy különálló fejezetben kerülnek ismertetésre.

7 Multicast Routing Protokoll

A multicast adatok továbbítása egy fa élei mentén történik. A fa gyökere lehet maga a forrás (SSM mód), pontosabban annak FHR-e, ekkor minden forráshoz külön multicast fa tartozik, illetve lehetnek a hálózatban dedikált routerek, melyek több forrás tartalmát is összegyűjtik (ASM mód), ekkor e dedikált források lesznek a gyökerek. E pontokat Rendez-vous Pontoknak (RP) nevezzük. Egy multicast csoportnak egy AS-en belül csak egy RP-je lehet, viszont egy pont több csoportnak is lehet dedikált pontja. A fa csomópontjait a multicast routerek alkotják, míg a leveleken a nyelők LHR-jei találhatók.

Ha egy LHR már ismeri a hálózatán lévő csoportokat, akkor valamilyen routing protokoll segítségével megpróbál a multicast fa tagjává válni.

7.1 A legjelentősebb routing protokollok

- DVMRP
- MOSPF
- BIDIR-PIM
- PIM-DM
- PIM-SM

7.2 Multicast Routing Protokollok legfontosabb tulajdonságai

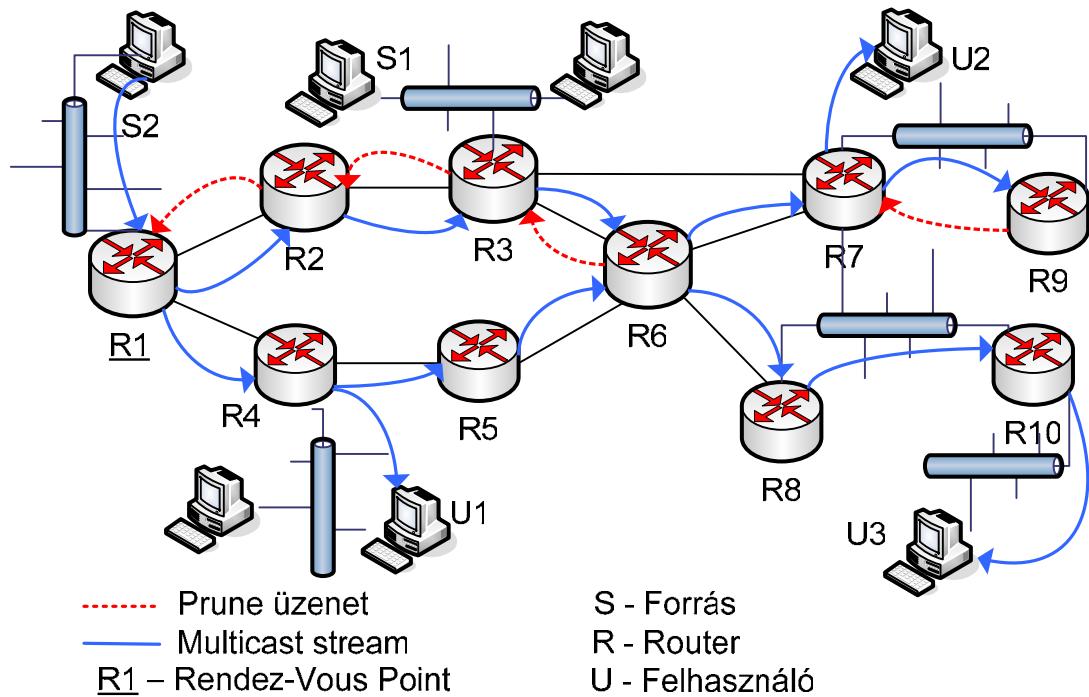
- „Sparse” vagy „Broadcast and Prune” működés
- Milyen multicast fatípust használnak

7.3 *Broadcast and Prune működés*

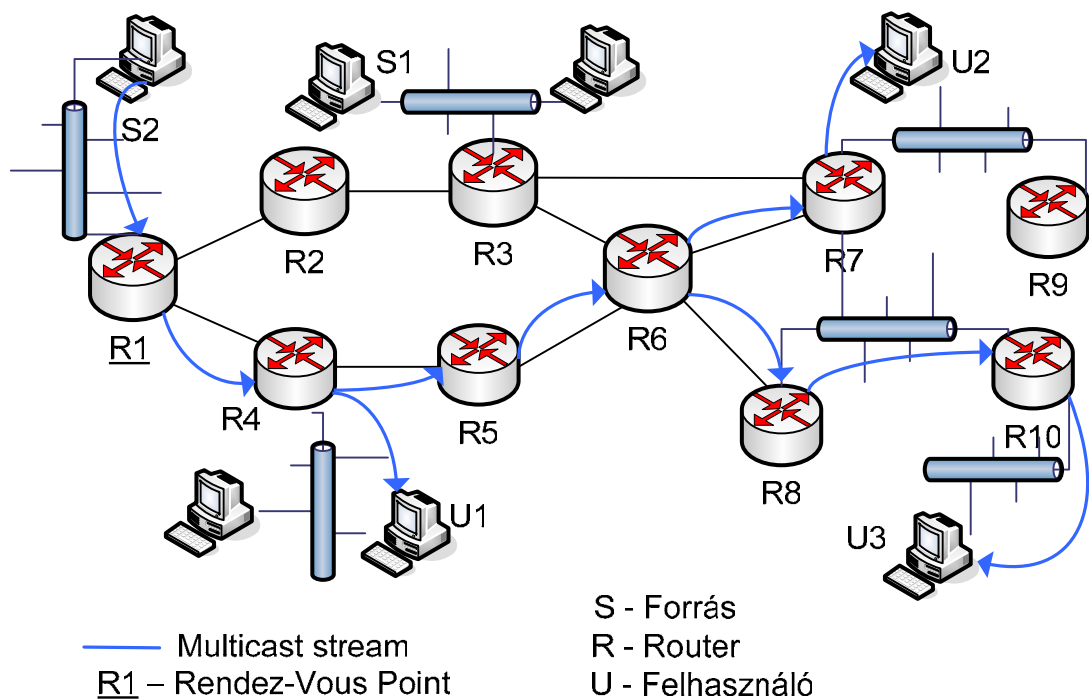
A Broadcast and Prune működés feltételezi, hogy nagy valószínűséggel minden routernek van legalább egy csoporttagja, így mindenkihez el kell juttatni a multicast adatokat. Ebből kifolyólag ezek a protokollok kezdetben olyan fát építenek, melyben valamennyi router szerepelni fog. Az adatok a gyökértől a fa élei mentén valamennyi routerhez eljutnak. Azok a útválasztók, melyeknek nincs szükségük a multicast adatokra, vagyis, akiknek nincs csoporttagjuk egy Prune üzenetet küldenek a hierarchiában fölöttük álló routernek, levágva magukat a multicast fáról. Ez az üzenet egészen addig felgyűrűzik, amíg nincs a routernek legalább egy gyermeke, aki igényelné a multicast streamet. Megelőzendő, hogy egy router örökre levágja magát a fáról, egy bizonyos periódus után a protokoll újraindul, így egy előzőleg levált router egy bizonyos idő után újra megkapja a streamet, s megint eldöntheti, hogy levállik a fáról vagy marad.

E csoportba tartoznak a PIM-DM, DVMRP protokollok.

A lenti ábra a protokoll kezdeti szakaszát ábrázolja. Az RP valamennyi kimenő interfészére továbbítja a multicast adatokat. Azok a routerek melyeknek nincs szüksége a streamre egy *Prune* üzenetet küldenek az upstream routernek (R9). Az R6 esete speciális. R6 két interfészen keresztül is megkapja a multicast streamet, az Reverse Path Check miatt az R3 felőli útvonal nem ideális, így erre küld egy *Prune* üzenetet. Mivel az R3-nal nincs csoporttagja, így tovább küldi a *Prune* üzenetet R1 felé, míg végül az egész útvonal leválasztódik. Az ábra feltételezi, hogy az U1,U2 és U3 felhasználók már IGMP protokollon jelezték, hogy csoporttaggá szeretnének válni.



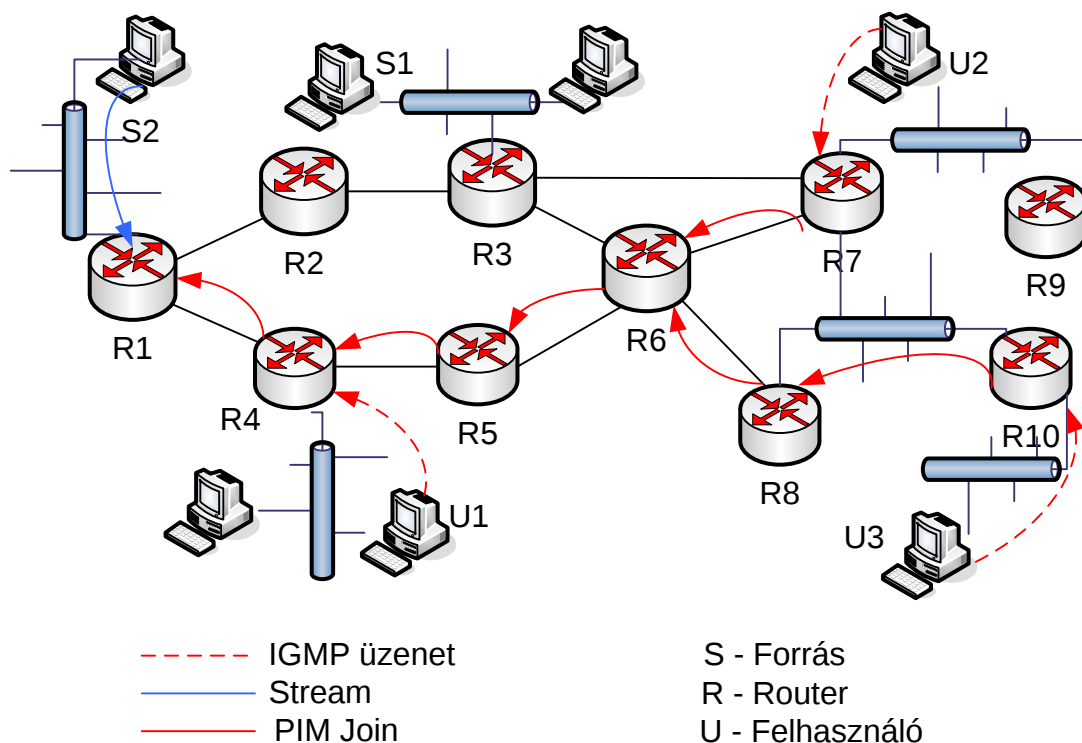
A következő ábra már a kialakult fastruktúrát mutatja be. A multicast stream a kialakult útvonal mentén jut el a feliratkozott csoporttagokhoz.



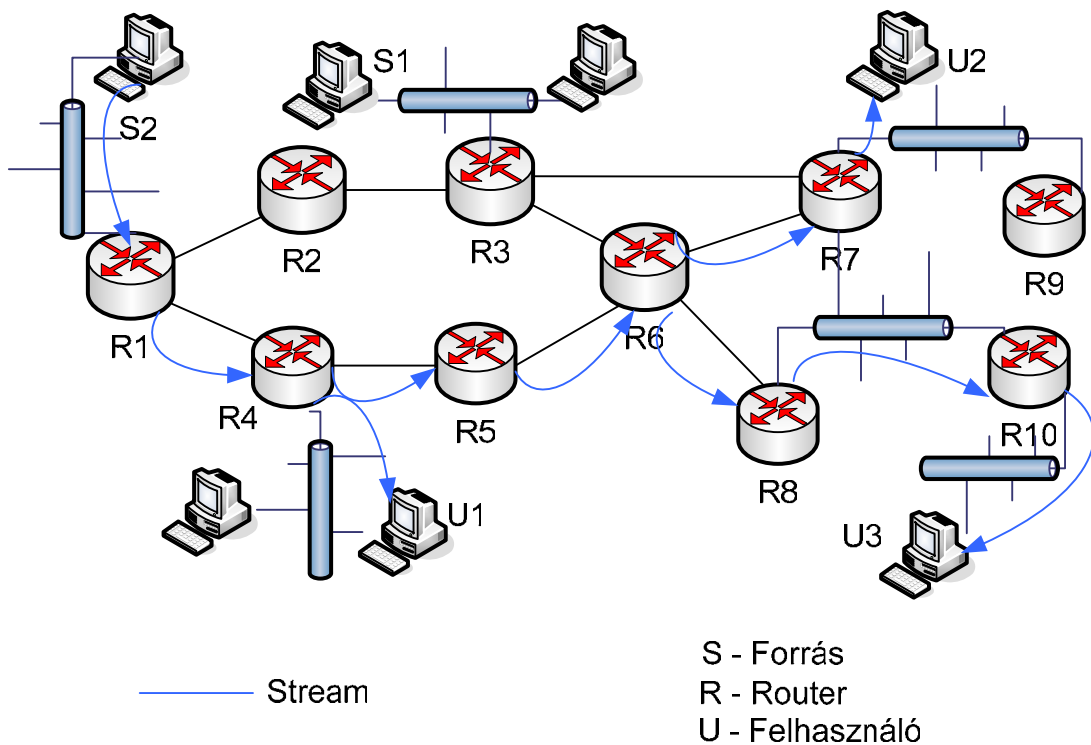
7.4 Sparse működés

A Sparse mód pont ellentétes a Broadcast and Prune működéssel. Itt az a kezdeti feltételezés húzódik az algoritmus mögött, hogy viszonylag kevés multicast felhasználó található a hálózaton. Miután a LHR-ek azonosították csoportjaikat, egy *Join* üzenetet küldenek a gyökér felé. Az út során valamennyi router, aki még nem kapta meg az adott streamet továbbküldi a *Join* üzenetet a gyökér felé, míg végül a kérés megérkezik egy olyan routerhez, mely már tagja a multicast fának, vagy a gyökértől kezdve kiépül egy új út. E csoportba tartoznak a BIDIR-PIM, PIM-SM, MOSPF protokollok

A lenti ábrán az U1, U2, U3 felhasználók IGMP protokollon jelzik multicast routereiknek, hogy csoporttaggá szeretnének válni. Ezek a routerek *Join* üzenetet küldenek az upstream routernek. Ha az upstream router még nem rendelkezik a streammel, akkor ő is küld egy *Join* üzenetet, mely végül vagy egy olyan routerhez jut, ahol már van stream, vagy eljut közvetlenül az RP-hez.



A következő ábra már a kialakult SHT-t mutatja. A multicast adatok a kialakult útvonal mentén jutnak el az aktív csoporttagokhoz.



7.5 Source Based Tree Protokollok

Ezen protokollok mindegyike forrásonként készít egy multicast fát. A fa gyökere a source FHR-je. A csatlakozni kívánó routerek a *Join* üzenetben egy (S_i, G) párost kell küldjenek, melyben megjelölik, hogy mely forrásoktól kívánnak adatokat fogadni. Ezen protokollok nagy előnye, hogy a multicast fa mindig az optimális útvonalakat tartalmazza. Ide tartoznak a PIM-SM, PIM-MD, MOSPF, DVMRP protokollok.

7.6 Shared Tree Protokollok

Az ide tartozó protokollok egyetlen fát építenek egy adott csoport számára. A fa gyökere az RP-on található. A forrásoktól a RP-hez a streamek a használt protokolltól függő eljárással jutnak el. Ez legtöbb esetben unicast küldést jelent, de például PIM-SM esetén ez RP-hez SPT-ken keresztül érkeznek meg a streamek. A csatlakozni kívánó routerek a *Join* üzenetben egy $(*, G)$ párost küldenek, így jelezve, hogy bármely forrástól érkező streamet meg szeretnék kapni. A protokoll előnye, hogy sok forrás esetén is jól skálázható, szemben az SBT

protokollokkal, azonban előfordulhat, hogy nem az optimális útvonalon továbbítódnak a multicast csomagok. SHT protokollok: PIM-SM, BIDIR-PIM

A PIM-SM protokoll mindkét működési módra képes, ezért található meg a SBT, illetve az SHT protokollok között.

7.7 Az upstream router meghatározása

Egy router az upstream interfészen keresztül küldi el a *Join* illetve *Prune* üzeneteket, valamint ezen az interfészen érkeznek a hasznos multicast adatok is. Más interfészekre érkező multicast adatok rend szerint eldobásra kerülnek, ezzel küszöbölve ki a duplikátumok létezését, illetve, hogy egy csomag körbe keringjen a routerek között.

Mindegyik routing protokoll az upstream router meghatározásához a reverse path forwarding ellenőrzést használja. Ez a mechanizmus a multicast fa gyökerének unicast címén alapszik. A mechanizmus lényege, hogy mikor egy multicast csomag érkezik, akkor a router megvizsgálja, hogy ő az adott interfészt használta volna adatok továbbítására az adott IP cím felé. Ha igen, akkor nagy valószínűséggel az adat az optimális útvonalon érkezett, és továbbítja kimenő interfészein a többi router felé. Ha nem, akkor a csomag eldobásra kerül.

A 90-es évek végén a multicast protokollok egy egészen új ága jelent meg. Ezek közös jellemzője, hogy függetlenek az AS-ben alkalmazott unicast útvonalválasztó algoritmustól, vagyis a multicast routing során a kész unicast routing táblákra alapozva járnak el. Két fajtája terjedt el, PIM-DM, PIM-SM. A PIM-DM-t működéséből fakadóan inkább kisméretű, sűrű csoportlétszámú AS-ben használják, míg a PIM-SM sokkal inkább alkalmas az internethez hasonló, nagy kiterjedésű, szétszórt felhasználójú hálózatokban.

8 Protocol Independent Multicast – Sparse Mode

Ez a protokoll, mint ahogy neve is jelöli független az alkalmazott unicast útvonalválasztó protokolltól, egyszerűen csak felhasználja az elkészült routing táblát.

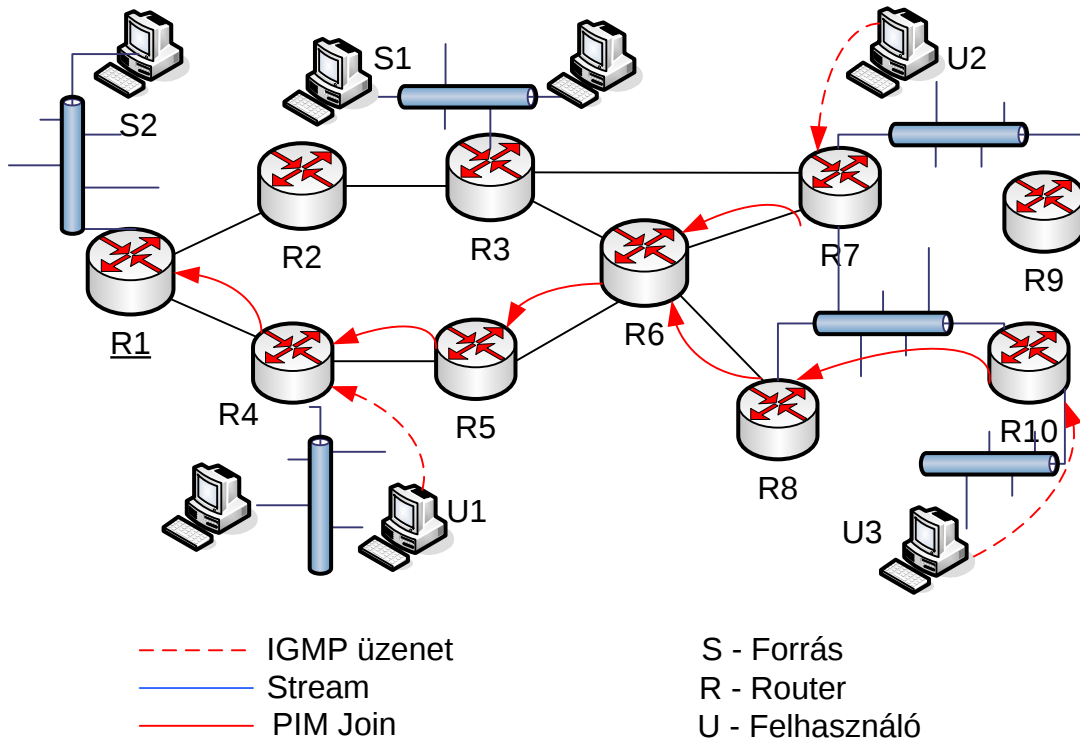
A sparse működésű protokollok családjába tartozik, és ez egyik legelterjedtebb ma használatos multicast protokoll.

A PIM-SM routerek periodikusan *Hello* üzeneteket generálnak, ezen üzenetek segítségével tartanak fent állapotmegőrző kapcsolatokat a szomszédos routerekkel. Miután felderítették a szomszédos multicast routereket, igény szerint *Join* illetve *Prune* üzeneteket küldenek az upstream router felé. A *Join* üzenet egy forrás és csoport párost tartalmaz (S_1, G_1). Amennyiben nincs specifikus forrás megadva (ASM), akkor a páros $(*, G_1)$ -re változik, ezzel jelölve, hogy valamennyi forrás streamjét meg szeretnék kapni. A *Join* üzenet addig halad felfelé az upstream routerek mentén, míg egy olyan routerhez nem ér, akinél megtalálható a kívánt stream, vagy eléri a multicast fa gyökerét.

A *Prune* üzenet hasonló elven működik, de ennek hatására megszűnik a stream forgalmazása a kérő router felé. *Prune* üzenetet akkor küldenek a routerek, ha már nincs aktív csoporttaguk.

Alapértelmezésben a PIM-SM SHT-t épít az állapottárolásból fakadó többletterhek minimalizálására. A SHT gyökere a Rendez-vous Point. Egy csoportnak egy AS-en belül mindig csak egy RP-je lehet, egy RP viszont több csoportnak is lehet dedikált pontja. Nyilvánvaló, hogy a protokoll egy sebezhető pontját jelenti az RP üzemképtelensége. Ennek elkerülésére egy automatikus RP választó mechanizmust használnak. Több lehetséges módszer is van, a legelterjedtebb az úgynevezett bootstep router használata. A routerek konfigurációja alapján csoportonként adott a lehetséges RP-k halmaza. Ezen routerek jelzik egy speciális entitásnak, a bootstrap routernek jelenlétüket. A bootstrap router valamilyen szempontok alapján kiválaszt egy RP-t és ezt kihirdeti a többi router felé is, így mindenki

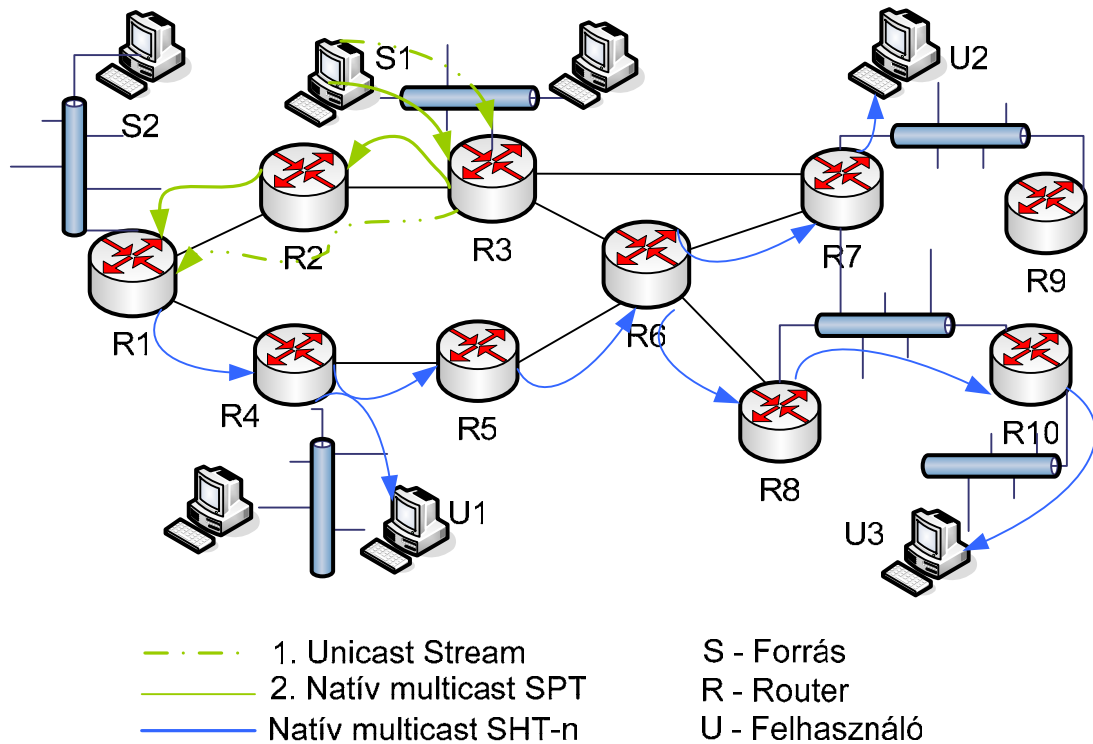
ismerni fogja az RP helyét. A bootstrap router kiválasztása is automatikusan, egy előre definiált candidate halmazból történik.



A fenti ábra az IGMP *Register* illetve PIM *Join* üzenetek útvonalát mutatja. A folyamat végére kiépül az SHT.

Ha kiválasztásra került egy csoport RP-je, a források a multicast streamet unicast *Register* üzenetekbe ágyazva kezdik el küldeni az RP felé. A küldés közvetlenül az RP-nek, unicast címe alapján történik. Mikor az RP megkapja az üzeneteket, a beágyazott multicast adatokat kibontja, majd direkt módon küldi tovább a multicast fán. Ez némi overheadet jelent a küldés folyamán, és nagy adatforgalom esetén jelentősen leterhelheti az RP-t. Ezt elkerülendő, mikor az RP megkapja a *Register* üzenetet, maga is egy *Join* (S,G) üzenetet küld vissza a forrás felé, melynek hatására csatlakozik az adott forrás SPT-hez. Ezeken a fákon tisztán multicast datagrammok formájában jut el a stream az RP-hez, és egyből tudja továbbítani az SHT-n. Az SPT-k kiépülése után minden forrástól két üzenet érkezik. Egy multicast csomag, illetve egy unicast csomagba ágyazott multicast üzenet. A felesleges unicast üzeneteket egy minden forrás számára periodikusan visszaküldött *Register Stop*

üzenettel szünteti meg az RP. Míg a *Register Stop* üzenetet visszaküldi, addig a forrás nem küld unicast csomagokat az RP-hez.



Az ábrán a source streamjének útvonala látható. Először unicast üzenetbe ágyazva jut el az RP-hez, majd az RP csatlakozik a forrás SPT-hez, és native multicast csomagokat kap.

Az SHT multicast fa viszonylag egyszerű módja a multicast adatok továbbításának, sajnos azonban az SHT útvonalai nem mindig optimálisak, így előfordulhat, hogy egy-egy forrás streamje feleslegesen sok erőforrást terhel le a hálózatban. Ezt elkerülendő a PIM-SM támogatja a SPT-k kialakítását is. Ha a LHR úgy dönt, hogy kedvezőbb lenne SPT-re váltani, mert optimálisabb az új útvonal, akkor egy forrás specifikus *Join* üzenetet (S_1, G) küld a forrás felé. Az SPT fa kiépítése teljesen megegyezik az SHT kiépítésével. Miután az új útvonal létrejött, az LHR egy speciális *Prune* (S_1, G, rpt) üzenetet küld a RP felé. Ennek az üzenetnek a hatására az első lehetséges routertől megszűnik a megadott source streamjének továbbítása az LHR felé, és a stream csak az újonnan kiépült, optimálisabb SPT-n keresztül jut el az LHR-hez.

A következő ábra azt mutatja, hogy az R6-os router úgy dönt, optimálisabb váltani az S1 SPT-re. Egy (S_1, G) forrásspecifikus *Join* üzenettel csatlakozik az SPT-hez. Miután az útvonal kiépül, egy (S_1, G, R_1) *Prune* üzenettel leválik a közös SHT-ről. Az üzenet az R4-ig gyűrűzik el, és addig megszűnik a régi útvonal. Az R6 továbbra is továbbítja multicast csomagokat az alatta lévő downstream routerek felé.

