

# IPv4 MULTICAST

## Alapprobléma

Multicast: többesküldés. Tipikusan egy forrás küld adatot sok receiver felé, pl. video v. audio folyamat. Elképzelhető hogy több forrás is van, pl. conference call.

Alapvetően a receiverek elviselik, hogy a stream "hibás", azaz kimáradnak belőle csomagok ill. egyes csomagok duplán érkeznek meg.

A hatékonyság kedvéért szeretnénk elkerülni külön-külön stream küldését minden egyes receiver felé. Ez egyrészt több erőforrást kíván a küldő oldalán is, másrészt megfelelően sok receiver esetén a küldő oldalán iszonyatos sávszélességigényű lenne. Ezért megpróbáljuk a folyamatot úgy küldeni, hogy csak akkor indítsunk új adatfolyamot, ha elkerülhetetlen.

Alapelemként az egy adott stream-ben érdekelt receiverek összességét (a címzettekét) Group -nak nevezzük. Egy adott group-ot a multicast célcím határoz meg. További alapelem a forrás (source), valamint a multicast tudással rendelkező router. Utóbbi végzi a hatékony csomagtovábbítást, amibe beletartozik a megfelelő pontokon az eddig egy irányból érkező csomagok több irányba szétszétlása (csomag többszörözés).

A multicast group tagjai a hálózatban szétszórva helyezkedhetnek el, bármely időpontban csatlakozhatnak a csoporthoz, illetve kiléphetnek belőle. Ilyen esetekben a megfelelő protollokoknak kell gondoskodniuk arról, hogy a felesleges irányokba ne menjen tovább, illetve a szükséges új irányok felé elinduljon a stream.

Az unicast routinghoz képest jóval bonyolultabb feladat a multicast routing, mivel a csomag továbbítása során nem lehet kizárólagos döntést hozni a célcím alapján. A multicast stream csomagjaiban a cél IP cím multicast cím, a küldő cím a forrás unicast címe.

A multicast adattovábbítás során a következő kérdésekre kell megoldást találni:

- Group management: új receiverek csatlakoztatása, kilépők eltávolítása.
- Intra AS routing: azonos felügyelet alatt levő routerek együttműködése az optimális csomagtovábbítás érdekében
- Inter AS routing: független rendszerek közötti multicast továbbítás
- L2 problémák: ethernet LAN-okban a last-hop router és az endpoint közötti problémák, esetlegesen a routerek közti ethernet szegmenseken fellépő problémák.

A multicast stream általában UDP csomagokat tartalmaz, de tartalmazhat RTP-t is. TCP-ről nyilván szó sem lehet, hiszen a TCP által megkövetelt újraküldések nem férnek össze a multicast jelleggel. A továbbiakban csak a stream továbbításának módjával foglalkozunk, a tartalommal nem.

## Multicast IP címzés

Multicast IP címzésre a class D osztályú címeket használjuk. Ezek a 224.0.0.0 és 239.255.255.255 közötti tartományban vannak. Ezek a címek mint célcímek szerepelnek, a multicast stream IP csomagjaiban a küldő címe a forrás unicast IP címe. A fenti tartomány több részre osztható. A főbb részek:

- Reserved link-local címek: 224.0.0.0 - 224.0.0.255

Csak egy adott linken (pl. ethernet szegmens) belül érvényesek. Küldéskor az ilyen csomagok TTL-je 1. A routerek nem továbbítják az ilyen csomagokat. Példa: 224.0.0.2 all-routers, 224.0.0.5 all-OSPF routers

- Limited scope 239.0.0.0-239.255.255.255 Organization-local ill. Site-local címzés (RFC2365) Ezen belül létezik az ún. relative address, ami az adott scope-hoz képest relatív címzést jelent. A relatív address azt jelenti, hogy az adott relatív cím a megfelelő scope legnagyobb címehez képesti offsetet jelzi (azaz pl. a 239.255.255.254 a local scope-n belüli 1-es relatív cím).
- GLOP 233.0.0.0/8 A középső két oktet az adott AS száma, így csak a <65536 AS-ek használhatják. Az adott AS az így generált GLOP blokkon belül szabadon rendelkezik a címekkel. (RFC3180)
- Global scope 224.0.1.0 - 238.255.255.255 IANA oszt, szervezetek közti multicast

A továbbiakat lásd az IANA multicast-addresses dokumentumában.

Érvényességi kör kijelölésére használatos még a TTL scoping, ahol a multicast stream TTL-jének értéke határozza meg a stream továbbításának "hatósugarát". Hasonlóan a normál IP-hez, a TTL értékét eggyel csökkenti minden router, és a 0 TTL-el érkező csomagokat eldobja. A TTL scoping hátránya többek közt az, hogy a hopcount nem feltétlenül ugyanakkora mindenfelé, így lehetséges, hogy a szükséges TTL már túlnyúlik a határon, így multicast forgalom szivároghat ki az adott területről.

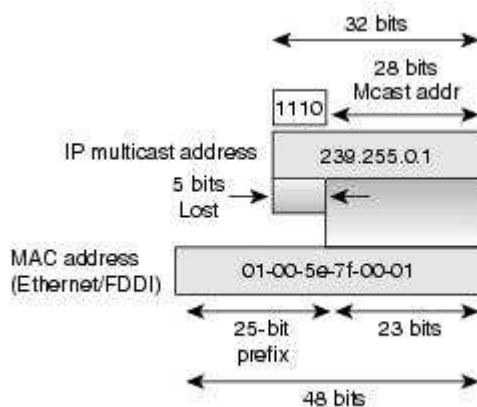
## Multicast MAC címek

Ethernet es FDDI hálózatokban leképezés kell az L3 multicast es L2 multicast címek között. Az ethernet cím 6 byte. Az első byte alsó bitje ha 1-ben van, az adott cím broadcast ill. multicast cím.

Az IPv4 címből leképezett multicast ethernet cím a következőképpen néz ki:

01:00:5e:0bbbbbbb:xx:xx

Az utolsó három byte megfelel a multicast IP cím utolsó 3 oktetjének, azzal a különbséggel, hogy az IP cím második oktetjének felső bitje elveszik. Ezzel összesen 5 bit veszik el az IP címből, így jőpár különböző IP group cím ugyanabba az L2 címbe képződik le!!! (például a 224.1.1.1, 224.129.1.1, ..., 239.1.1.1, 239.129.1.1) Ennek történelmi okai vannak.



Token ring esetén az eljárás más, ezzel nem foglalkozunk.

## Group management

A receiver hostok egy adott csoporthoz való csatlakozásukat az "Internet Group Management Protocol" segítségével végzik. Az IGMP protokoll egy adott host es a first-hop router közti kommunikációt írja le.

Az IGMP v1 (RFC 1112) alapvető üzenetei a "Host Membership Query" es a "Host Membership Report".

Az adott hálózathoz (tipikusan ethernet szegmensre gondoljuk) tartozó router periodikusan Query üzeneteket küld az all-hosts (224.0.0.1) címre (TTL=1, tipikusan 60 sec?? az intervallum). Az üzenet természetesen tartalmazza a lekérdezett group címét. A group tagjai erre válaszolnak egy Report üzenettel. A report üzenetet egy member host nem rögtön a query vétele után küldi, hanem egy random timer lejártá után (10 sec max IGMP v1-ben). A report üzenet címzettje a group. Ha a timer lejártá előtt a host vesz egy report üzenetet, akkor ő már nem küld újabb reportot az adott groupra vonatkozóan. Ez a report suppression mechanizmus a felesleges forgalom megszüntetését próbálja megoldani, hiszen a routernek ha csak egy receiver van akkor már kell küldenie a multicast stream-et. Ha az adott időn belül nem érkezik report válasz, akkor a router nem küld tovább.

A grouphoz csatlakozáskor egy host un. unsolicited report üzenetet küld. A csoport elhagyása "csendben" történik, azaz a csoportot elhagyó host nem küld semmilyen üzenetet. Ha mindenki elhagyta a csoportot, akkor a router nem fog report választ kapni a következő query üzenetre, így legközelebb csak akkor veszi észre, hogy már nem kell a streamet küldenie.

Az IGMPv2 (RFC2236) az IGMPv1 néhány fogyatékosságát igyekszik kiküszöbölni. Itt kerül bevezetésre a "Group Specific Query", ami az adott csoport címére megy (szemben az IGMPv1-nél az all-hosts címre küldött query-vel). Emellett a "Generic Query" is megjelenik, itt a lekérdezett csoport címe 0.0.0.0 .

Itt kell megemlíteni, hogy az IGMPv1-ben nem volt külön támogatás arra az esetre, amikor egy szegmensben több IGMP képes router is van. Az IGMPv2-ben ennek feloldására egy "IGMP Querier Election" került bevezetésre. Először minden router küld query-t, majd a legalacsonyabb IP számú router lesz a Querier router, a többiek nem küldenek query-t. Ez a szerep tehát különválik a "Designated Router" szereptől!

Az IGMPv2-vel kiküldött query-ben megjelent egy új mezo, a "Maximum Response Time". Ez alapértelmezésben 10 sec (az IGMPv1 fix ideje). A hostoknak ezen időn belül kell választ küldeniük a query-re. A query suppression mechanizmus ugyanúgy működik mint az IGMPv1-ben (figyelembe veve a kiküldött maximum response timer-t). A query küldés ideje itt is 60 másodperc alapértelmezésben.

A csoport elhagyásának módja változott az IGMPv2-ben. Egy host a csoport elhagyásakor egy Leave üzenetet küld, amire válaszul a router kiküld egy group specific query-t. A csoport aktív tagjai a normál mechanizmussal válaszolnak. Ha nem maradt az adott csoportban aktív tag, akkor a megadott időn belül nem érkezik report, így a router abbahagyhatja a stream küldését. Így hamarabb kiderül, hogy aktív-e a csoport (IGMPv1 esetén rossz esetben egy query periodus+10sec után derül csak ki).

Ha egy adott szegmensben IGMPv1 es v2 tudású eszközök is vannak, akkor a következő szabályok érvényesek:

- ha egy host IGMPv2-t tud, de a router IGMPv1-et tud, akkor a query-re v1-es verziójú reportot kell küldenie. Továbbá a csoport elhagyásakor nem kötelező Leave üzenetet küldenie, de ha küld, az nem gond.
- ha a query-küldő router IGMPv2-t tud, akkor detektálnia kell hogy van v1 host a hálózatban, mivel a v1 hostok nem tudnak a query response interval-lal mit kezdeni. Ebben az esetben a routernek ignorálnia kell a Leave üzeneteket, hiszen az erre válaszul küldendő group specific query-t nem fogják felismerni a v1 hostok.
- Ha több router van egy hálózatban, akkor minden routernek ugyanazon IGMP verziót kell futtatnia. Cisco routereknél ezt kézzel kell konfigurálni.

Az IGMPv3 (RFC 3376) az elozoekhez képest jóval bonyolultabb. A legfőbb különbség, hogy a host kijelölheti, hogy a multicast stream forrásai közül melyik forrást akarja venni ill nem venni. Ezzel lehetővé válik az ún. Source Specific Multicast. A források megválasztásának lehetősége új API

bevezetését is megköveteli, így a meglévő multicast applikációkat újra kell írni.

A Query-Response mechanizmus is változott az IGMPv2-hoz képest. Az IGMPv3 routerek ezentúl az all-igmpv3-routers (224.0.0.22) címen hallgatnak. Ezt a címet a hostok nem figyelik. Ebből következően további fontos változás, hogy eltűnik a response suppression mechanizmus, azaz minden host válaszol egy vett Query üzenetre. A küldendő válaszokat a hostok random idő múlva küldik, de ez az idő az előzőekhez képest jóval hosszabb is lehet (a max. response time mező speciális értelmezésű).

A források kiválasztásával szükségessé vált a General es Group-Specific Query üzeneten kívül egy Group-and-Source-Specific Query. Ezzel a router az adott csoporton belüli források iránti igényeket kérdezheti le.

Az IGMPv3 megköveteli, hogy a router a v1 es v2 -t is ismerje.

## L2 multicast problémák

Ethernet hálózatokban tipikusan switcheket használunk. A switch alapvetően minden portjára floodolja a multicast célcímmű ethernet frame-eket, mivel nem tud róluk semmit sem (hiszen ezek a címek sosem jelennek meg mint forráscímek). Ez azonban nem kívánt jelenség, hiszen így olyan portokon is megjelenik adat, ahol nem kellene (es a switch erőforrásait is negatívan befolyásolhatja ez).

Néhány mechanizmus létezik ennek megoldására: IGMP snooping, CGMP.

IGMP snooping esetén az L2 eszköz lehallgatja a multicast csomagokat, és az így meghallott IGMP join ill. leave üzenetek alapján frissíti a forwarding tábláját. Kellemetlen ebben az, hogy ehhez a stream minden csomagjába bele kell néznie, ez jócskán igénybe veheti az erőforrásait.

CGMP: Cisco Group Management Protocol. Cisco proprietary protokoll. Működéséhez szükséges hogy az L2 es L3 eszközök is ismerjék. Mikor egy host küld egy IGMP join üzenetet a routernek, akkor a router e csomag alapján elküldi az L2 eszközöknek a receiver unicast source címét es a multicast group címét. Ez alapján az L2 eszköz el tudja dönteni hogy melyik portjára forwardolja ezt a streamet.

Hasonló eszköz a GMRP: GARP Multicast Registration Protocol (IEEE 802.1P) GMRP tudás a hostban és a L2 switch(ek)ben szükséges. Itt a host (tipikusan az IGMP join-nal egy időben) küld egy GMRP join üzenetet a switchnek. Így a switch a multicast forgalmat a megfelelő portra tudja korlátozni. A switch periodikusan GMRP query-eket küld a hostoknak. A hostok erre válaszolva megerősítik multicast receiver státuszukat. A csoport elhagyásakor a host küldhet leave üzenetet, vagy egyszerűen nem válaszol a GMRP query üzenetekre.

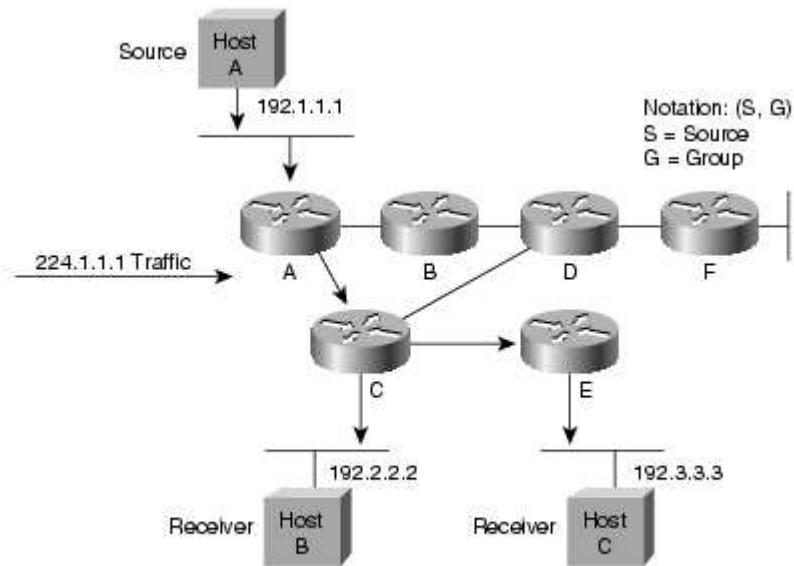
IGMP processzorigény!!!!

RGMP!!!! IGMP proxy???

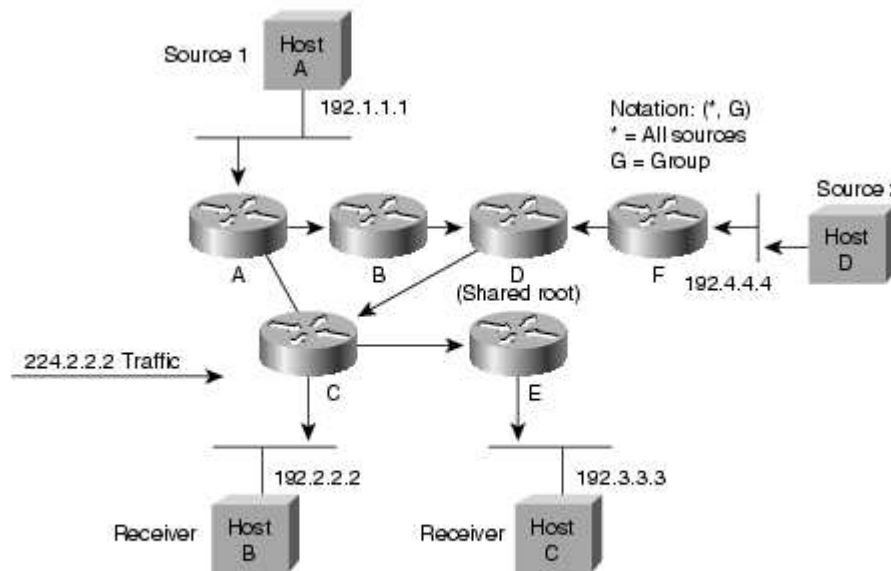
## L3 intraAS (intradomain) routing

### Alapfogalmak

A multicast adatfolyam útja a forrástól a receiverekig fa alakú (distribution tree). Egyik fajtája a Shortest Path Tree. Ez a fa a forrás es a vevők közti minimális költségű feszítőfa. A fa gyökere a forrás, ezért különböző források különböző SPT-eket határoznak meg (ezért nevezik Source Tree-nek is). Jelölése (S,G).



A distribution tree másik fajtája a Shared Tree. Ebben az esetben több multicast folyamat a hálózatban egy közös pontig halad valamilyen úton, majd a közös ponttól az összes stream egy megosztott fán (shared tree) halad a receiverekig. A közös pontot a PIM "Rendezvous Point" -nak nevezi.



A SPT tehát összességében több memória felhasználását igényli, viszont optimális utat ad a forrás és receiver között. A Shared Tree viszont kevesebb memóriát igényel, de extra késleltetést vezethet be, mivel nem az optimális (de elég jó) utat jelöli ki.

A Distribution Tree felépítését az egyes multicast routing protokollok különböző módon végzik.

A csomagtovábbítás multicast esetben jelentősen eltér az unicast esettől. Unicast esetben a router a döntést a csomagok célcíme alapján hozza meg. Multicast esetben a célcím a csoport címe, így a döntést a forrás címe alapján kell meghozni. A csomagtovábbítás során a routerek az ún. "Reverse Path Forwarding"-ot használják. Ez azt jelenti, hogy a router egy multicast csomagot csak akkor továbbít, ha az a "forrás felé eső" interfészen jött, azaz a distribution tree mentén jutott el a routerig. Azok a csomagok, amelyeknél az RPF teszt nem igaz, eldobásra kerülnek. Természetesen a csomagot a router a forrás irányába sohasem küldi vissza! A "forrás felé eső" interface megállapítása a multicast routing protokoll információja alapján történik. **RESZLETEZ, ÁBRA**

Az intra-domain multicast routing protokollokat két csoportra oszthatjuk: dense es sparse módú protokollokra.

Dense módban a routing protokoll feltételezi, hogy mindenhol vannak vevők, így a routerek flood-olják a forgalmat. Azok a routerek, amelyek feleslegesen kapják a forgalmat (tudják hogy nincs receiverük) "prune" üzeneteket küldenek. Így a fa felesleges ágai sorban visszametszésre kerülnek. Mindez periodikusan megismétlődik. Ebből látható, hogy ilyen esetben nem optimális a sávszélesség felhasználása, mivel a multicast forgalom periodikusan olyan helyekre is eljut, ahol arra nincs igény.

Sparse módban a fa csak olyan irányba épül ki, ahol van igény a multicast stream vételére (explicit join behavior). A mai hálózatokban gyakorlatilag csak sparse módú protokollt használnak multicast routing-hoz.

Dense módú protokollok a DVMRP, (MOSPF,) PIM-DM. Sparse módú a PIM-SM es CBT (Core Based Trees, RFC2201). Utóbbival nem foglalkozunk, mivel nem terjedt el. A többi Dense Mode protokollt is csak a routing sajátosságainak megmutatása érdekében tárgyaljuk.

## **DVMRP - Distance Vector Multicast Routing Protocol**

A DVMRPv1-et az RFC 1075 definiálta, és ma már elavult. Jelenlegi verziója internet draft. A DVMRPv3 szinten draft.

A DVMRP az MBONE routing protokollja volt, de ma már kevésbé használt. Dense módú, azaz induláskor eláraszt. A nem kívánt irányokban levő routerek prune üzenetet küldenek az upstream routernek. A prune operáció egy idő után "timeout"-ol, így újra elárasztás történik.

A RIP-hez hasonló protokoll. A routerek subnetelérhetőségi információkat cserélnek ki egymással, mint a RIP-ben, de itt a subnetekhez tartozik subnet mask is (RIPv1-ben ilyen nincs). A RIP-hez hasonlóan itt is van "végtelen" metrika, ennek értéke 32. A DVMRP a Distribution Tree-t az ún. "Truncated Broadcast Tree" kialakításával építi fel. A "Truncated Broadcast Tree" definíciója: Az S1 subnet TBT-je az az S1 gyökerű, minimális úthosszú feszítőfa, ami a hálózatban levő összes routert érinti. Ha egy router több irányban is eléri az adott subnetet, akkor a kisebb IP számú router felé építi ki a fát. A DVMRP routing tábla az összes subnet TBT-jét tartalmazza.

A DVMRP a RIP minden hibáját hozza magával. Nem skálázható. Ma már nem használjuk multicast hálózat routingjához.

## **MOSPF Multicast extension to OSPF (RFC1584)**

Az elterjedt OSPF protokollban az unicast Link State Advertisement mellett az ún Group Membership LSA-t használja. Az unicast LSA-hoz hasonlóan terjednek ezek is a hálózatban (intra-area és inter-area). A Dijkstra algoritmust használja a SPT létrehozásához. Ezek alapján hozza létre a routerben a forwardoláshoz szükséges állapotokat. Minden egyes forrás-csoport pároshoz külön számol topológiát.

Mivel az OSPF kiegészítése, ezért csak OSPF alapú hálózatokban működik. Skálázási problémák miatt nem használatos, ui. minden egyes forráshoz külön futtatja a Dijkstra algoritmust, ami jelentős erőforrásigényt jelent. Például nem stabil hálózati linkeknél, illetve ki-be lépő tagoknál komoly hatással lehet a routerekre a folyamatos fa-újraszámolás miatt. Ami még rosszabb, az adminisztrátornak nincs lehetősége ezt a viselkedést befolyásolni!

Mivel minden belépő és kilépő receiver ill. forrás esetén kiépül egy SPT, ezért itt nincs szükség flooding-ra.

## **PIM-DM Protocol Independent Multicast, Dense Mode (RFC 3973)**

Dense módú protokoll. A "Protocol Independent" elnevezést onnan kapta, hogy nem önmaga épít fel routing táblázatokat, hanem a meglévő unicast routing táblát használja. Szintén flood+prune módon működik, de a DVMRP-tól eltérően nem egy előre felépített fa mentén teszi ezt (hiszen ilyen nincs neki), hanem minden olyan interface-n floodol, ahol van egy PIM-DM szomszéd, vagy a group egy tagja közvetlenül elérhető. A routerek prune üzeneteket küldenek, ha a multicast stream rossz irányból jön (a non-RPF interface-n), illetve az RPF interface-n akkor, ha nincs downstream irányban érdekelt fél (router v. group member). Ez a viselkedés alapértelmezésben 3 percenként megismétlődik.

Lényeges elem még az ún. Assert mechanizmus. Ezt abban az esetben használják a routerek, amikor egy adott irány két v. több routeren át is elérhető (pl. redundancia). Ebben az esetben a routerek Assert üzenetet küldenek egymásnak, amiben a "distance" és "metric" értékek szerepelnek. Ez alapján az alacsonyabb értékét elérő router (akinek jobb útja van a forráshoz) nyer, döntetlen esetben a nagyobb IP szám a nyerő.

A flood/prune ill. az Assert mechanizmus azonban problémákat is okozhat (pl. route loop stb.). Emiatt és a dense mód kis hatékonysága miatt csak kis hálózatokban ajánlott használni.

## **PIM-SM Protocol Independent Multicast, Sparse Mode (RFC 4601)**

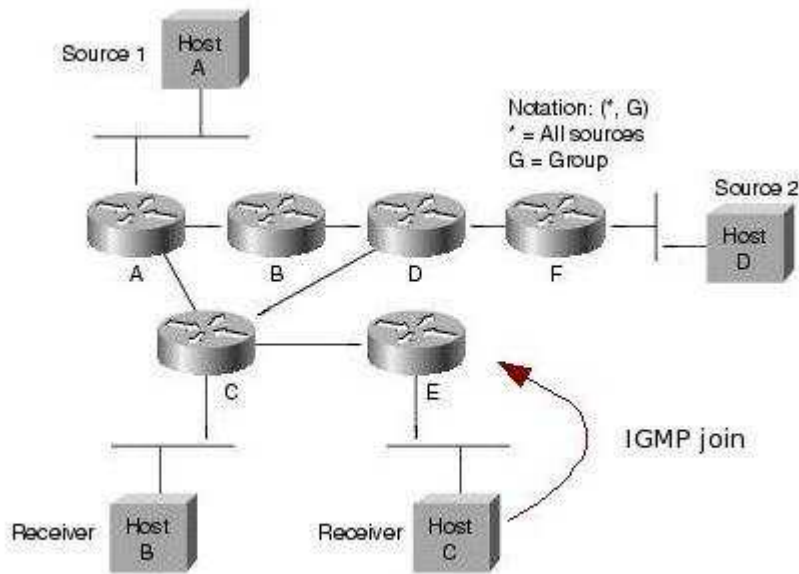
A PIM-SM-ben alapvető fogalom a Rendezvous Point. Ez egy olyan PIM router, ami kitüntetett szerepet tölt be, segítve a források és receiverek egymásra találását. A multicast forgalom az RP-től a receiverekig kialakított Shared Tree mentén folyik (ezt hívják RPT Rendezvous Point Tree -nek). Az egyes receiverek az IGMP mechanizmusával csatlakoznak egy csoporthoz. Ennek hatására a PIM router (leaf router ill. last-hop router) explicit módon csatlakozik az RP-hez, amely folyamat során kialakul a shared tree az RP felé.

A források regisztrálják magukat az RP-hez. Ez úgy történik, hogy a forrás első csomagjai hatására a first hop PIM router ezeket a csomagokat becsomagolja egy speciális Register csomagba, és unicast módon továbbítja azokat az RP felé. Az RP kicsomagolja ezeket a csomagokat, és továbbítja a tree-n a receiverek felé. Emellett Join üzenetekkel elkezd kiépíteni egy SPT -t a forrás felé. A tree felépülése után a folyam a tree mentén halad az RP felé, ekkor az RP Register stop üzenettel jelzi a forrásnak, hogy abba hagyhatja a Register csomagok küldését. Az SPT kiépülése és a Register stop üzenet feldolgozása közt eltelt időben a multicast forgalom két példányban halad a hálózatban (egyszer a Register üzeneteken belül, egyszer pedig az SPT mentén).

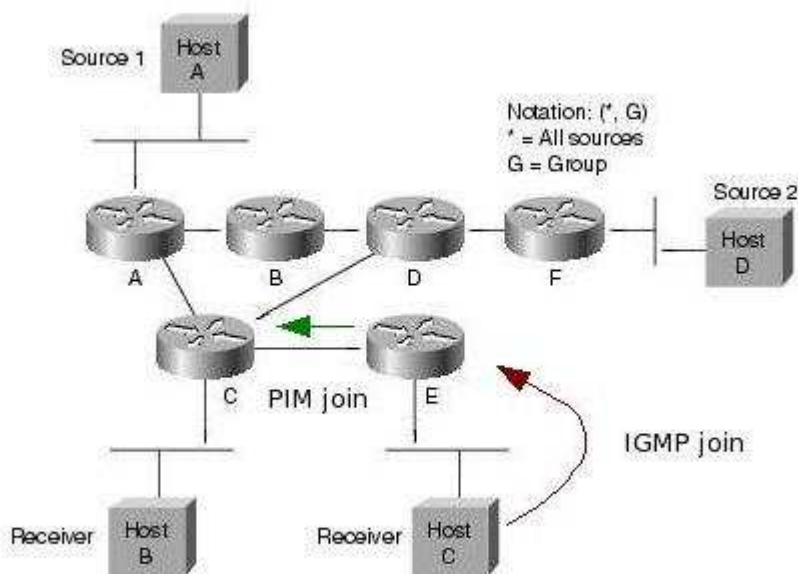
Ezután a folyam a forrástól az RP-ig, majd onnan az RPT-n a receiverekig halad. Ez nem biztos hogy optimális (extra késleltetés, nem minden esetben optimális útvonal, RP mint szűk keresztmetszet). Ezért ilyenkor lehetőség van arra, hogy a folyam továbbítása átálljon a forrás-receiver SPT-re (SPT switchover). Ekkor a last-hop router (S,G) join üzeneteket küld a source felé, így kiépítve az SPT-t. Az SPT átállás általában az adott group "traffic rate" küszöbértékének elérése után kezdődik meg. A küszöbértéket kb/s-ben értelmezzük, és az adott csoport RPT-n folyó forgalmának sebességét adja meg (total aggregate rate, azaz nem pillanatnyi rate). A router másodpercenként számolja ki ezt az értéket. A küszöbérték meghaladásakor a csoport minden forrása felé elkezd kiépíteni az SPT-t. Az RFC4601 szerint a küszöbérték implementációtól függő (a last hop router konfigurációjától függ). Cisco routereknél ez rögtön megtörténik, mert a küszöbérték 0. Lehetséges azonban végtelen küszöb is, aminek hatására sohasem történik SPT átállás. Az átállás után a receiver felé egyrészt az SPT-n másrészt az RPT-n is megjön a forrástól származó csomag, ezért ekkor az RPT mentén (S,G) prune üzeneteket is küldenek a routerek. A forrástól az RP irányába továbbra is folyik adatforgalom, hiszen lehet hogy más receiverek nem állnak át az SPT-re. Ez csak akkor szűnik meg, ha az RP az összes interface irányából kapott (S,G)

prune üzeneteket az RPT-re.

Ha az összforgalom az adott küszöbérték alá csökken, akkor a last-hop router a "shared tree switchback" algoritmus segítségével visszaállíthatja a forgalmat a shared tree-re. A küszöbérték ellenrzése percenként történik. A switchback során az SPT lebomlik, és join üzenetekkel újra felépül a shared tree.

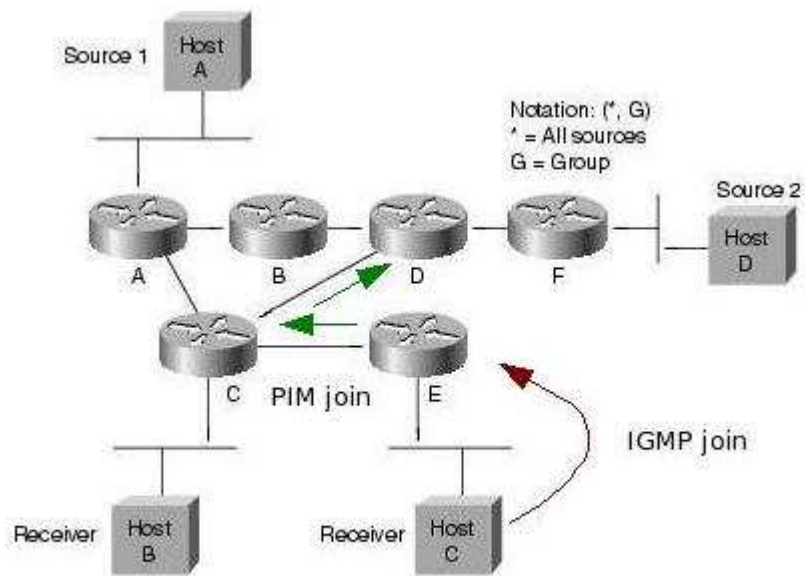


1. Egy receiver csatlakozik egy csoporthoz: küld egy IGMP join üzenetet a last-hop routernek.

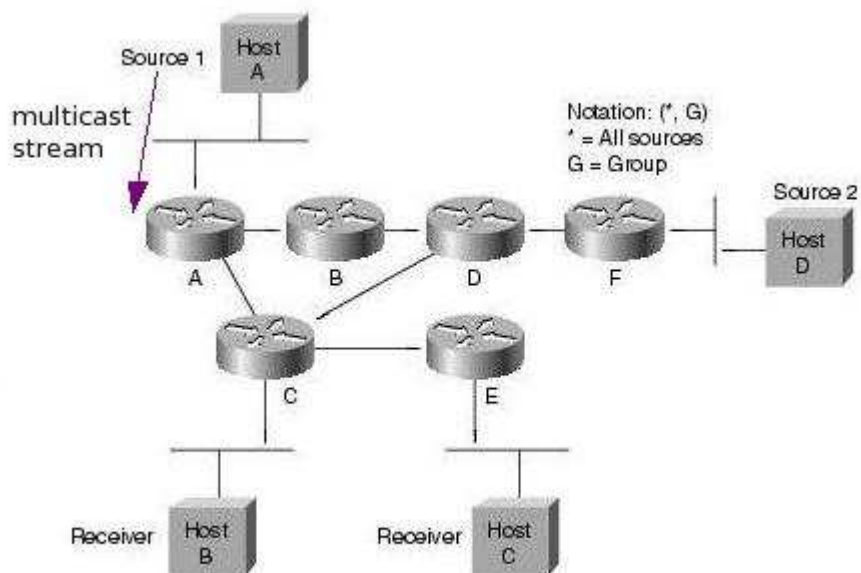


2. A router PIM Join üzenettel elkezd kiépíteni a shared tree-t az RP felé. Az E routerben létrejön a (\*,G) állapot.

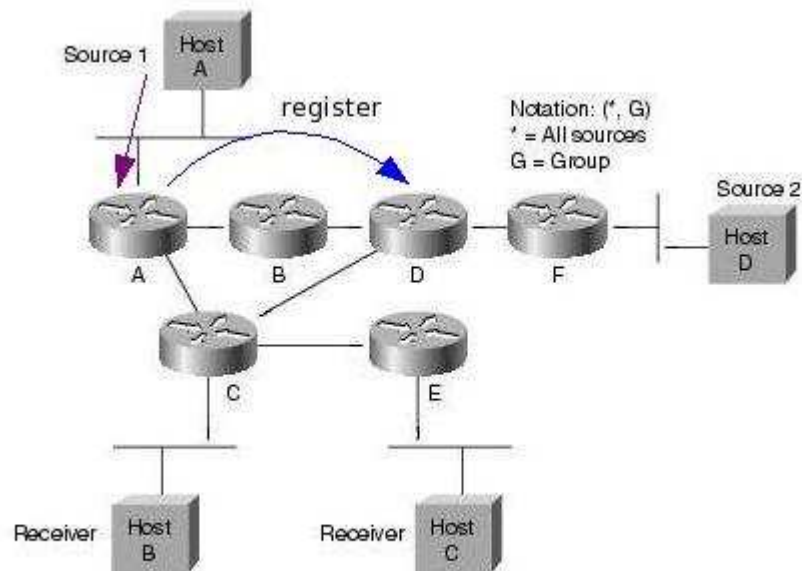




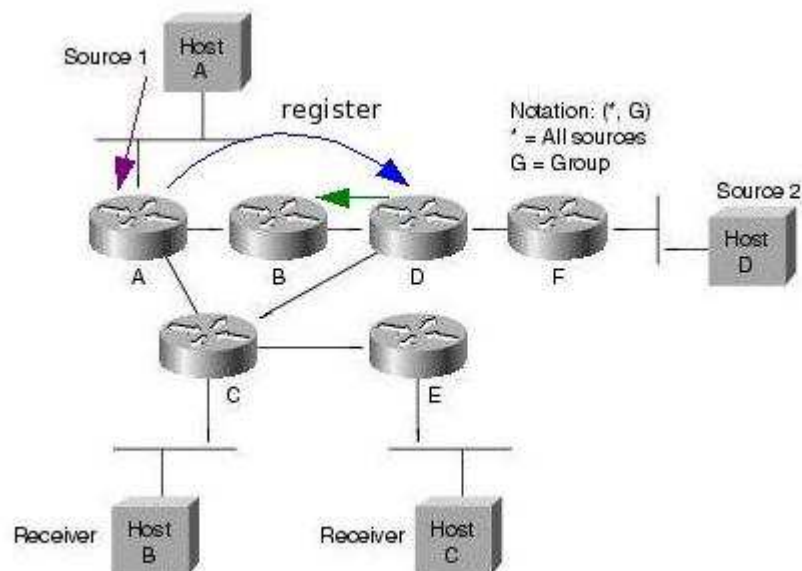
3. A fa kiépüléséig a közbenső routerek PIM Join üzeneteket küldenek az RP felé, és mindenhol létrejön a (\*,G) állapot.



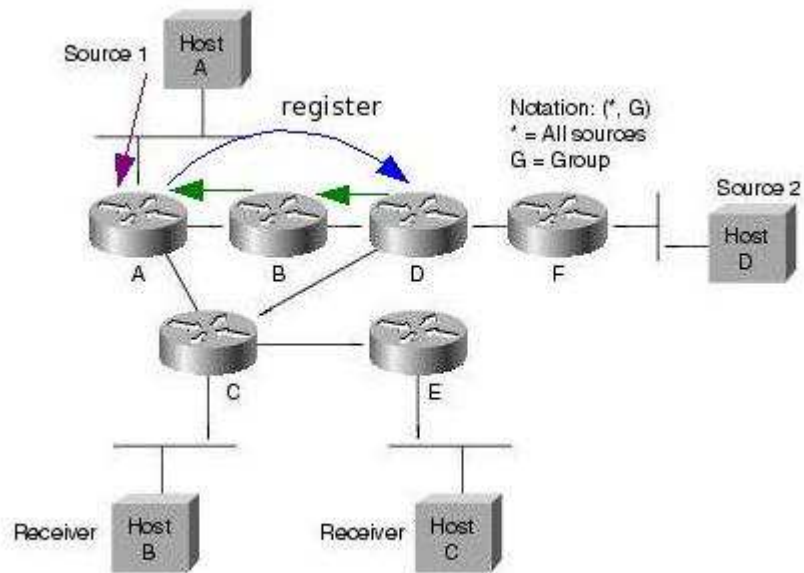
4. Egy forrás elkezd az adott csoportban adni.



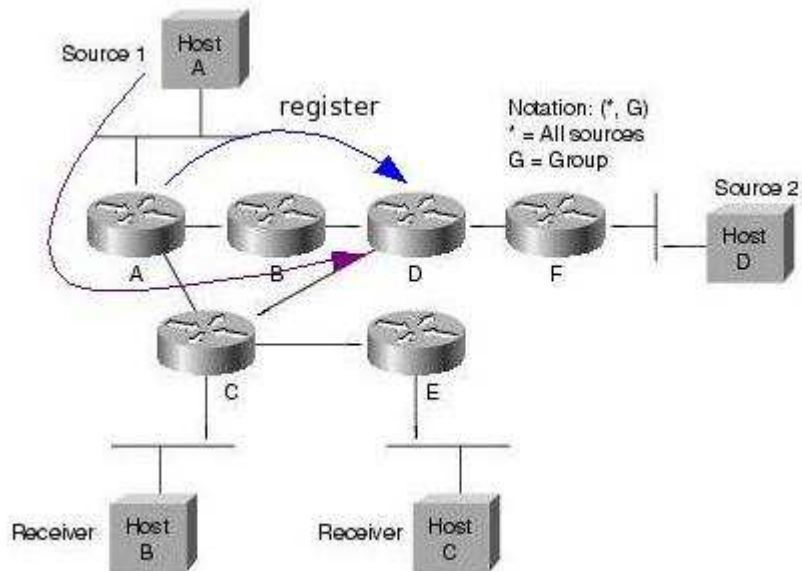
5. Ennek hatására a first-hop router register üzenetet küld az RP-nek. A register üzenet unicast csomag, tartalmazza a multicast forrás csomagjait.



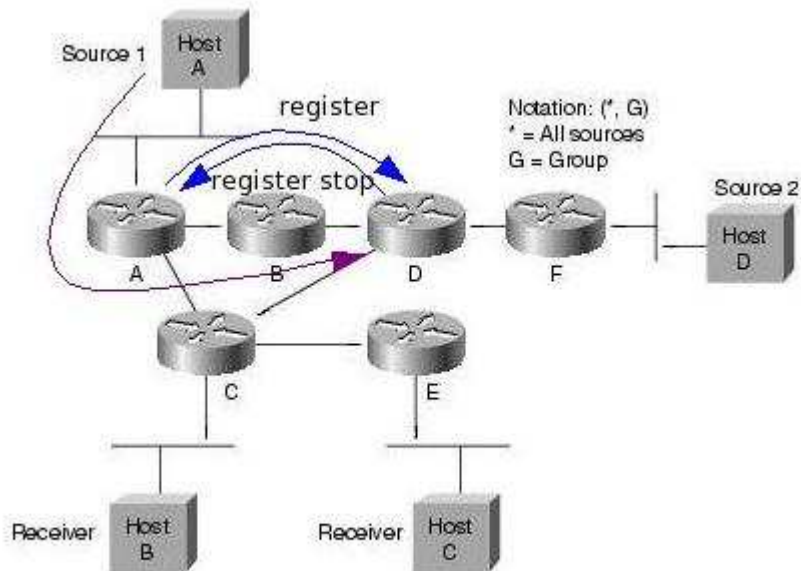
6. A register üzenet hatására az RP PIM join üzenetekkel kiépíti az SPT-t a forrás felé. A közbenső routerekben létrejön az (S,G) állapot. Emellett az RP az esetleg létező RPT-n továbbítja a register-ből kicsomagolt multicast stream-et.



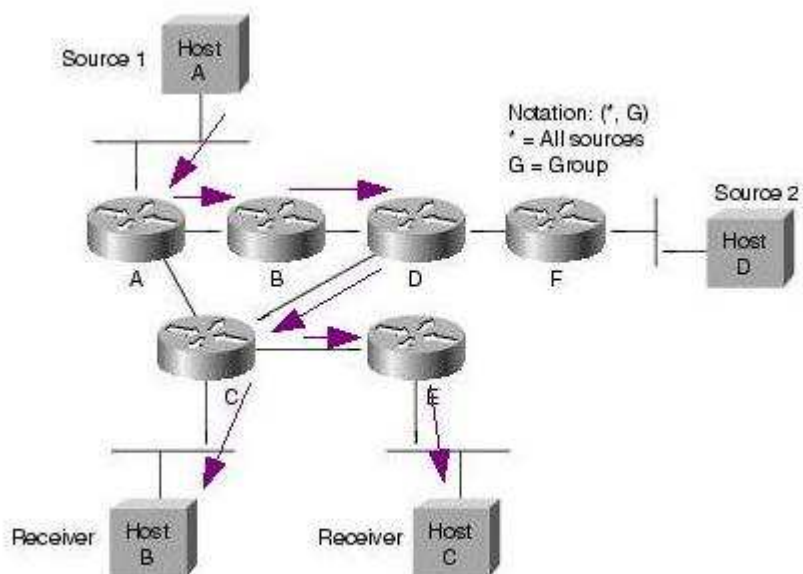
7. Képzül az SPT a forrás és a first-hop router közt. Az SPT-n mentén a routerekben létrejön az (S,G) állapot.



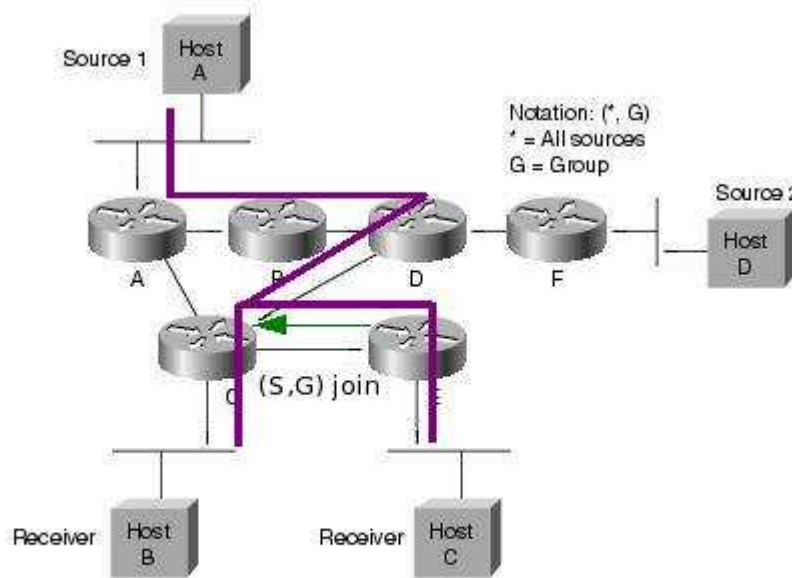
8. A multicast adatfolyam az SPT mentén is halad az RP felé.



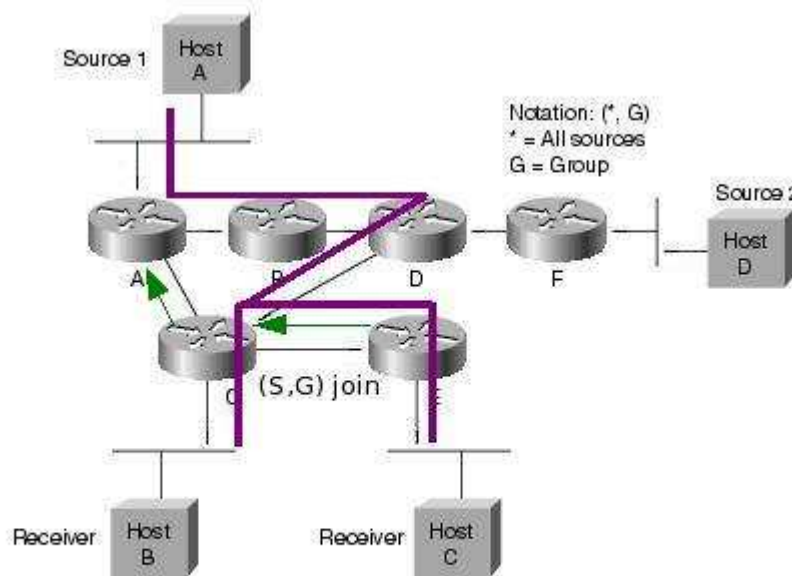
9. A natív multicast folyam megjelenésekor az RP register stop üzenetet küld a first-hop routernek, ez a register-hez hasonlóan szintén unicast csomag.



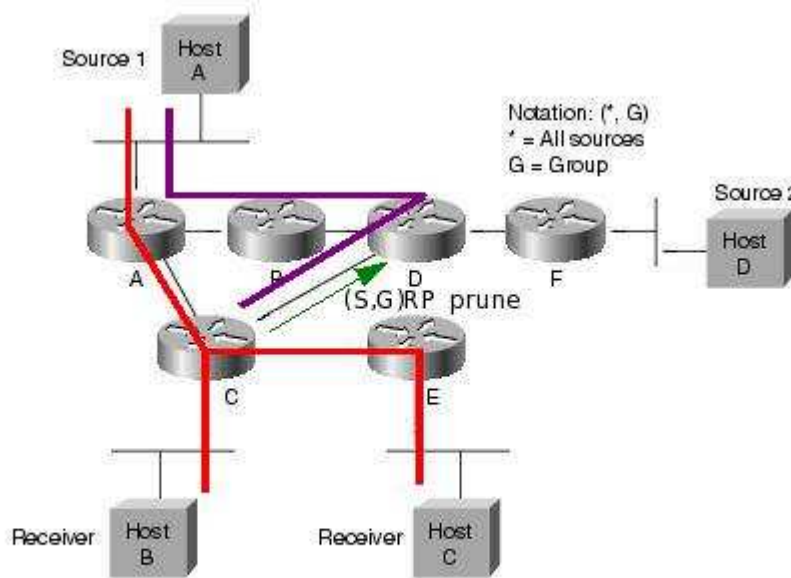
10. A multicast folyam a forrás-RP-shared tree mentén folyik.



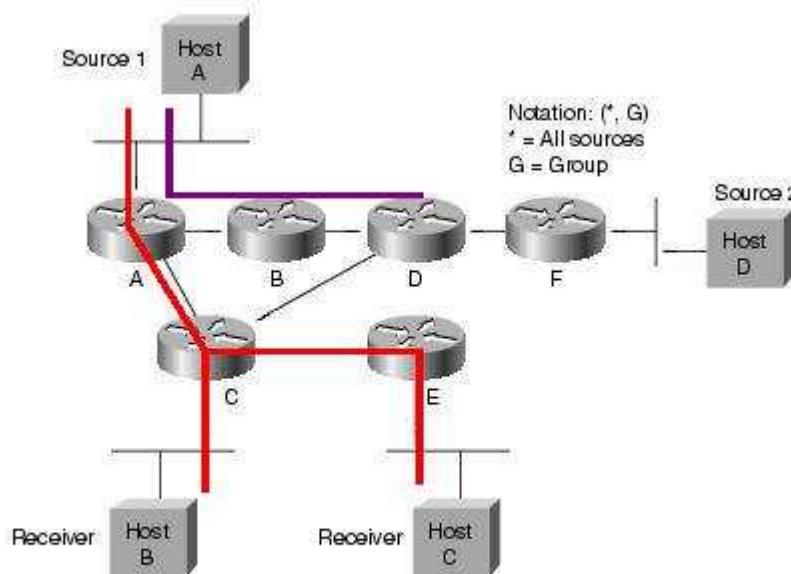
11. A megadott küszöbérték meghaladásakor az **E** router megindítja az SPT switchover folyamatot: (S,G) join-t küld a source irányába.



12. A közbenső routerek további (S,G) join-ok küldésével kiépítik az SPT-t.



13. Az SPT felépülése után C egy (S,G)RP prune üzenetet küld, mert nincs züksége az RPT-n érkező forgalomra.



14. A teljes forgalom az SPT-n folyik. Amennyiben az RPT az RP-ig vissza lett bontva, az RP további (S,G) prune üzenetekkel lebontja a source felé az SPT-t.

A Rendezvous Point router szerep tehát egy kitüntetett routeré. Az RP kijelölésére több módszer is lehetséges.

Statikus RP esetén minden hálózati router konfigurációjába kézzel bekerül az RP címe. Ez nyilván nem minden esetben hatékony ill. rugalmas. Két automatikus mechanizmus létezik az RP kijelölésére: a Cisco Auto-RP funkciója, és a PIMv2 Bootstrap Router mechanizmusa.

Az auto-RP mechanizmusban a candidate RP-k a "Cisco Announce" címre multicastolják információjukat. Ezt az ún Mapping Agent routerek veszik. A nagyobb IP számú RP-t választják, és ezt az információt küldik tovább a "Cisco Discovery" multicast csoportba (chicken-egg, ezek a csoportok dense módban vannak!!). Az összes cisco router veszi a discovery csoportot, és eltárolja a csoport-RP összerendelést.

A PIMv2 BSR mechanizmus során kiválasztódik egy Bootstrap Router, BSR (valami egyszerű eljárással).

A candidate RP-k a BSR felé küldik információjukat (unicast). A BSR kiválasztja a csoportokhoz az RP-t, és ezt az információt hirdeti. Az egyes routerek az all-pim-routers csoport használatával hop-by-hop flood-olják az RP információt. A BSR mechanizmus és az admin scoping nem működik együtt, mert a candidate RP-k üzenetei átmehetnek a multicast határokon.

Itt megemlíthető még az "Anycast RP" is (RFC 3446), aminél az RP címe anycast cím. Ebben az esetben egy adott (vagy több) csoporthoz több RP is tartozhat. A last-hop routerek az RP anycast címéhez képest építik ki az RPT-t, ami a valóságban az anycast megoldás miatt a topológiai közel eső RP irányába kiépített fát fogja jelenteni. Az egy anycast csoportba tartozó RP-k MSDP (ld. később) segítségével szereznek tudomást a csoportokhoz tartozó forrásokról (fontos megjegyezni hogy az MSDP peeringhez NEM az anycast címet használják). Az RP-k ezután SPT-t építenek ki a megismert források irányába.

Az "Anycast RP" egy másik mechanizmusát az RFC 4610 írja le. Itt az anycast RP-k közt nincs MSDP peering. A belépő forrás Register üzenetét a hozzá "tartozó" anycast RP unicast segítségével juttatja el a többi, előre konfigurált anycast RP-hez.

További redundanciát megvalósító eljárás a Neighbor discovery. Ez arra jó, hogy shared (pl. ethernet) szegmenseken több PIM router közül Designated routert jelöljünk ki. Minden szóbjövő router periodikusan PIM Hello üzeneteket küld az "all-pim routers" címre. A nagyobbik IP számú nyer, ő fogja ezután a join/prune stb. dolgokat végezni. Ha a DR kiesik, akkor egy bizonyos timeout letelte után új sorsolással új DR választódik.

## **Bidirectional PIM**

A kiépített fa kétirányú, és minden forgalom az RP gyökerű fán megy (nincs SPT). IETF draft (draft-ietf-pim-bidir-08.txt 2005. oct) Támogatás???

## **PIM Source-Specific Extensions (RFC3569)**

A receiverek "csatorna"-ként kezelik a (S,G) párost, ahol S ismert forrás (külön úton megszerzett információ). A vevők IGMPv3 segítségével eleve a kiválasztott source által adott streamet igénylik. Így nincs szükség a shared tree infrastruktúrára. A megfelelő IGMPv3 üzenetek hatására a last hop router eleve a kiválasztott S source SPT-jét építi fel ill. ahhoz csatlakozik. Ezzel mintegy "megoldódik" a hagyományos multicast egyik problémája, ami a kéretlen források kezelése, hiszen eleve az adott source SPT-je épül ki (megjegyezzük, hogy a Cisco implementációjában van lehetőség az RP-ben az ip pim accept-register <acl> utasítással annak szabályozására, hogy kitől fogad el az RP Register csomagokat). Egyszerűsödik továbbá a multicast group címek kiosztásának problémája is, mert ezentúl csak az adott source-n belül kell kezelni a csoportcímek ütközését (SSM-re külön tartományok).

## **L3 inter-AS (inter-domain) multicast routing**

Nem minden AS tud multicast routingot. Ennek a problémának egyik kikerülése, hogy a multicast AS-ek közt tunnel segítségével juttatjuk el a multicast streamet a címzett AS-hez. A másik lehetőség: BGP kiegészítés (MBGP, RFC2283), ami segítségével protokoll családonként küldhetünk útvonalválasztási információt. Ezek alapján kialakítható a multicast topológia, e fölött a multicast topológia fölött tudja a PIM-SM felépíteni a fákat. Kiemelendő, hogy az MBGP segítségével csak a multicasthoz használható topológiára derül fény; ahhoz hogy valóban továbbítódjon egy multicast folyam, továbbra is szükséges hogy a közbenső routerek PIM-et tudjanak.

További kérdés, hogy az egyes RP-k hogyan szereznek tudomást a saját AS-ükön kívüli forrásokról (egy aktív forrás csak a saját AS-én belüli RP-hez regisztrál). Erre való az MSDP, Multicast Source Discovery

Protocol (RFC3618). A szomszédos RP-k egymást tájékoztatják a saját AS-ükön belüli forrásokról, és ezt az információt továbbadják. Így minden RP tud mások forrásairól is, és szükség esetén kiépítheti felé az SPT-t (MBGP-ből ismert topológia alapján). SSM esetén nem kell. (Cisco MSDP SA-filter: lehetőség hogy szűrjük a local (private&loopback) forrásokat, domain-local forrásokat, scoped forrásokat, auto-RP forrásokat es SSM forrásokat).

## További kérdések

LOAD BALANCING+PIM????

Amiről nem volt szó: multicast forrás elhelyezése

Reliable Multicast: PGM Pragmatic General Multicast RFC3208

RFC 4410 Selectively Reliable Multicast Protocol (SRMP)

multicast+MPLS

multicast+MPLS VPN (mVPN)

---

### Források:

- [http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito\\_doc/ipmulti.htm](http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/ipmulti.htm)
- <http://www.multicasttech.com/faq/>
- RFC-k, IANA dokumentumok
- <http://www.javvin.com/protocolGMRP.html>
- Jákó A. Multicast Routing, Networkshop 2001 <http://splash.eik.bme.hu/>
- <http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t2/dtbipim.htm>