

Segédlet az IP Multicast működésének megértéséhez

Bevezetés

Napjainkban egyre több helyen használnak olyan együttműködést támogató szoftvereket, technológiákat, amelyek lehetővé teszik, hogy egyszerre kettőnél több fél kapcsolódjon egymáshoz és együttesen végezzen munkát, kommunikáljon vagy valamilyen módon információkat osszon meg egymással. Ahhoz, hogy ez hatékonyan megvalósulhasson, szükséges a kommunikációs infrastruktúra felkészítése az esetlegesen megnövekedő igények megfelelő szintű kiszolgálására. Ennek egyik lehetséges módja az IP multicast technológia alkalmazása. A következőkben röviden bemutatásra kerül az IP multicast technológia alapvető működése.

IP Multicast

Az IP hálózatokban három csomagtovábbítási módszert különböztetünk meg, úgymint unicast, multicast és broadcast.

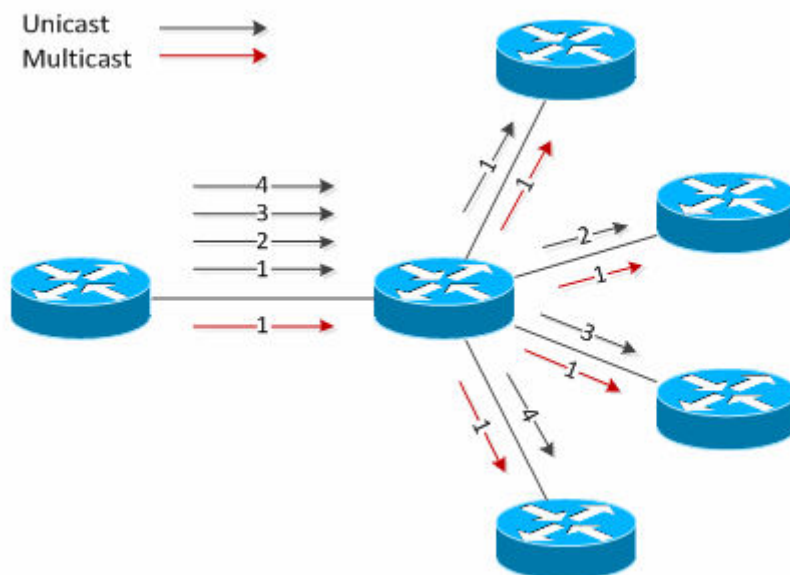
- **Unicast** továbbítás esetén mind a forrás, mind pedig a nyelő egyetlen entitás, így az adatsomagok kiindulási és végpontja egyedi.
- **Broadcast** csomagtovábbítás esetében a forrás egyetlen entitás, de a nyelő az adott szórási tartomány összes további végpontja.
- **Multicast** továbbítás logikájában a két módszer közé helyezhető, melyben mind a forrás, mind pedig a nyelő lehet egy vagy több entitás is, vagyis az összes végpont egy csoportja.

Multicast csoport alatt végpontok olyan összességét értjük, amelyek ugyanazon adatfolyamra tartanak igényt. Ahhoz, hogy egy végpont megkapja ezeket az adatfolyamokat szükséges, hogy csatlakozzon ahhoz a csoporthoz, melyhez tartozó forgalmat meg szeretné kapni. Az ilyen végpontokra egy speciális D típusú IP-címmel hivatkozunk, melyre küldött adatokat a csoport összes regisztrált tagja megkapja. Ahogy a unicast címeknél, úgy a D típusú IP címek esetén is beszélhetünk külön fenntartott címtartományokról, melyeket az IANA határozott meg:

- **Link-local címek:** (224.0.0.0 – 224.0.0.255) hálózati protokollok által használt címek, melyeket a forgalomirányítók (továbbiakban routerek) nem továbbítanak.
- **Egyéb fenntartott címek:** (224.0.1.0 – 224.0.1.255) protokollok és alkalmazások számára kiosztott címek.
- **Globális címek:** (224.0.2.0 – 238.225.225.225) IANA kezelése alatt álló globális címek.
- **Lokális címek:** (239.0.0.0 – 239.255.255.255) autonóm rendszeren belül szabadon kiosztható címek.

A multicast csomagok ethernet hálózaton történő továbbításához a adatkapcsolati réteg szintjén is kijelöltek egy címtartományt. Ez a címtartomány a 01:00:5E:xx:xx:xx tartomány, mely a címben 23 szabad bitet tart fent, amire az IP-címeket leképezhetjük. Mivel a D típusú IP címek első 4 bite minden esetben 1110 ezért csak 28 bitet kell leképezni a fent említett 23 bitre, így egy multicast MAC címre egyszerre 32 multicast IP címet képezünk le. Ez a sajátosság teljesítmény problémákat okozhat a végpontok, valamint egyes L2 eszközök esetében is.

A multicast előnye, hogy a forrás nem küld el minden csomagot annyi példányban, ahány cél van — mint azt unicast továbbításnál tenné — hanem csak egyetlen példányt küld el, amelyet ha szükséges, a hálózati eszközök sokszorosítanak, annyi példányban ahány interfészükön azt tovább kell küldeniük.



1. ábra – Unicast és Multicast továbbítás

A multicast adattovábbítás működéséhez mindenképpen szükség van valamilyen eljárásra, mely lehetővé teszi a különböző csoportok tagsági viszonyainak nyilvántartását, valamint a különböző végpontok elérését segítő útválasztási mechanizmust. Az előbbit a csoportkezelés, míg az utóbbit a multicast útválasztási protokollok valósítják meg.¹

Csoportkezelés

A multicast adattovábbítás egyik alappillére a csoportkezelés, aminek a feladata a különböző csoportok valamint azok tagjainak nyilvántartása. Ez nem azt jelenti, hogy rendelkezni kell egy központi nyilvántartással az összes csoportról illetve azok tagjairól, hanem azt, hogy az infrastruktúrát alkotó hálózati eszközöknek tudniuk kell, ha egy interfészükön keresztül elérhető valamely csoport legalább egy tagja, akár közvetlenül akár valamely távolabbi szegmensben.

A csoportkezelés eszköze egy vezérlő protokoll, az Internet Group Management Protocol (továbbiakban IGMP), melynek segítségével a hálózaton található routerek értesülnek a velük megegyező LAN szegmensben lévő végpontok multicast csoportokhoz kötődő állapotváltozásairól. Az így nyert információkból ezután képesek egy olyan nyilvántartás kiépítésére, amely tartalmazza, hogy melyik interfészeiken milyen csoportokból vannak regisztrált tagok.

IGMPv1

Keretek felépítése

Az IGMP első verzióját még sok hiányosság jellemzi, úgymint a csatlakozás és a kilépés

¹ Jeff Doyle: Routing TCP/IP, Volume II, Cisco Press, 2001 420-424 p.

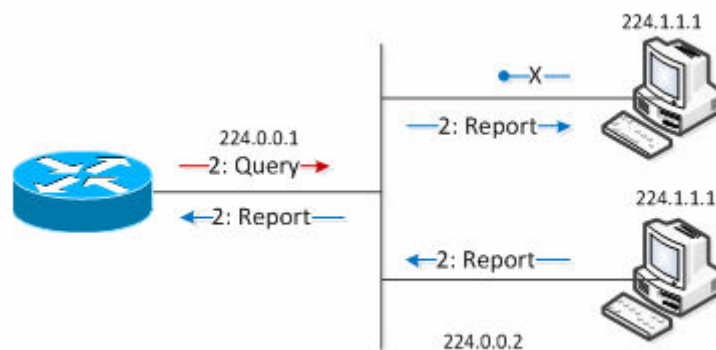
készletetése, valamint az útválasztó protokollra való támaszkodás a kérdező router kiválasztásához. Az IGMP üzenetek IP csomagokba ágyazott 64-bites üzenetek, melyek négy mezőt tartalmaznak, úgymint verzió (4 bit), típus (4 bit), ellenőrző összeg (16 bit) valamint a csoport cím (32 bit).

Verzió (4)	Típus (4)	Nem használt (8)	Ellenőrző összeg (16)
Csoport cím (32)			

Működés

Lekérdezés

A csoportok állapotának fenntartása egy kérdés-felelet mechanizmussal történik, mely két üzenet típus segítségével valósul meg. Ez a két üzenettípus a Membership Query és a Membership Report. A kérdező router megadott időközönként ismétlődve egy Membership Query üzenetet küld az adott LAN szegmensen az összes multicast képes végpontot megcímző 224.0.0.1 címre. Ezt követően a végpontok egy 0 és 10 másodperc közötti idő elteltével, Multicast Report üzenettel válaszolnak azokra a csoport címekre, amelyeknek tagjai. Mivel a válaszok az adott csoport címére érkeznek ezért azt minden csoporttag megkapja. A routernek elég csoportonként egyetlen válasz — mert csak azt kell tudnia, hogy van-e az adott interfészen csoporttag — ezért az első választ megkapva a többi tag már nem küld választ a router Membership Query üzenetére.



2. ábra - Lekérdezés menete

Csatlakozás

Amikor egy végpont csatlakozik egy csoporthoz, nem szükséges a következő Multicast Query megvárása, hanem egy úgynevezett váratlan Multicast Report üzenettel jelzi a router számára tagságát.

Kilépés

A csoport elhagyására a protokollnak ez a verziója nem rendelkezik eljárással. A router nem értesül külön a végpontok csoportelhagyásairól, kizárólag arról, ha már nem található tag a LAN szegmensen. Ez úgy történik, hogy nem kap választ a Multicast Query üzeneteire, aminek hatására három próbálkozás után kiveszi az adott multicast csoport forgalmát az interfészre továbbítandó adatok közül. Mivel a Multicast Query üzenetek között eltelt idő alapbeállítások szerint hatvan másodperc, ez akár hárompercnyi fölösleges forgalmat is jelenthet az adott interfészen.²

² Beau Williamson: Developing IP Multicast Networks, Volume I, Cisco Press, 1999, 56-62 p.

IGMPv2

Keretek felépítése

A kérdés-felelet mechanizmus ebben a verzióban négy különböző IGMP üzenettípus segítségével valósul meg. A lekérdezést továbbra is a Query üzenet biztosítja, amit azonban két altípusra bontunk. Ezek az általános illetve a csoport-specifikus Query üzenetek. A válaszadásra továbbra is a Report üzenet szolgál, mely lehet v1 illetve v2 report üzenet is. A negyedik üzenettípus a csoportból való kilépés jelzésére szolgáló Leave Group üzenet. Az IGMPv2 üzenetek felépítésüket tekintve is eltérnek a korábbi verzió üzeneteitől. Az üzenetek tartalmazznak egy 8 bites típus mezőt, melyben a típuskódok a korábbi verzióval való kompatibilitás miatt a korábbi verzió és típus mezők uniójaként jön létre, egy 8 bites „Maximum Response Time” mezőt, egy 16 bites ellenőrző összeget valamint egy 32 bites csoport címet.

Típus (8)	Max. Válaszidő (8)	Ellenőrző összeg (16)
Csoport cím (32)		

A 8 bites „Max Response Time” mező arra szolgál, hogy a lekérdező router megadhassa azt a maximális időt, amit a válaszoló végpontok, a válasz küldése előtti véletlen várakozási idő számításánál felhasználnak. Ez az érték az egyes verzióban statikusan 10 másodperc volt.³

Működés

Lekérdezés

A kérdés-felelet mechanizmus működése megegyezik a korábbi verzióban használt eljárással, különbséget csupán a két lekérdező altípus használata jelent. Az általános lekérdező üzenetekben a csoport cím csupa 0 értéket tartalmaz, míg csoport specifikus lekérdezésnél a mező az adott csoport címét tartalmazza. Az általános lekérdezés a 224.0.0.1-es minden multicast képes végpontot címző multicast címre, míg a csoport specifikus lekérdezést az adott csoport címére küldi a router. A válaszadási késedelem csökkentésére a csoport-specifikus lekérdezésben a Max Response Time kisebb, mint az általános lekérdezésben (alapbeállítások szerint 1 másodperc).

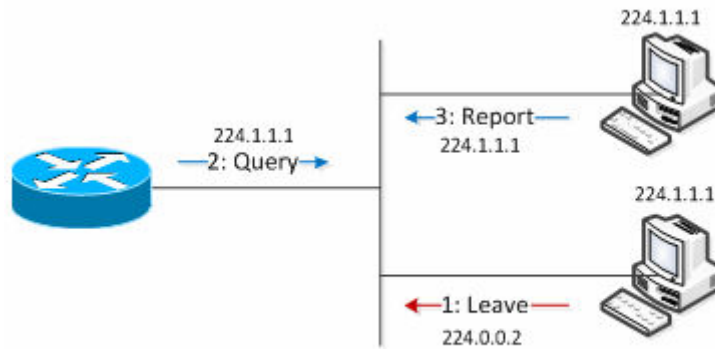
Csatlakozás

A végpontok csoportokhoz csatlakozását továbbra is a végpont által küldött váratlan Report üzenet jelzi.

Kilépés

A csoport elhagyásával kapcsolatos hiányosságokat, késedelmeket a Leave Group üzenet bevezetésével küszöbölték ki. Amikor egy végpont elhagy egy csoportot egy ilyen üzenetet küld a 224.0.0.2, „összes router”-t címző speciális multicast címre. A kiválasztott lekérdező router ennek hatására egy csoport specifikus lekérdezést küld a LAN szegmensre, hogy megtudja, maradt-e még a csoporthoz tartozó tag az adott interfészen. Amennyiben az elküldött lekérdezésre nem érkezik válasz, egy „Last Member Query”-nek nevezett általában 1 másodperces időtartam után még egy lekérdezést küld, amelynek sikertelensége esetén kiveszi az adott csoportot az interfészre továbbítandó multicast forgalom közül.

³ W. Fenner: Internet Group Management Protocol, Version 2 (RFC 2236)



3. ábra - Végpont kilépése

Az IGMPv1-el ellentétben a kettes verzió már biztosít eljárást a lekérdező router kiválasztására abban az esetben, ha az adott LAN szegmensen több router is található. Amikor egy router aktív állapotba kerül, elküld egy általános lekérdező üzenetet, melyet minden router megkap. Az üzenetben megtalálható a router forrás címe, amit a routerek a saját címekkel összehasonlítva eldöntik, hogy lekérdező routerek-e. Végeredményként a LAN szegmensen a legalacsonyabb IP címmel rendelkező router lesz a lekérdező router.⁴

Kompatibilitás

Az IGMP második verziója kompatibilis az első verzióval, azonban annak egyszerűbb eljárásai miatt, valamint amiatt, hogy az IGMPv1 eszközök nem ismerik fel az IGMPv2 üzeneteket, a különböző verziók együttműködése nehézkes. Abban az esetben, ha a kommunikáló router IGMPv2-t használ, de a végpont csak v1-et, akkor a router nem használhat csoport specifikus lekérdezéseket a csoporttagság eldöntésére és nem várhatja el, hogy a végpontok üzenettel jelezzék, ha kilépnek egy csoportból.

IGMPv3

Az IGMP protokoll harmadik verziójának legnagyobb újítása az úgynevezett Source Specific Multicast technológia támogatása. A SSM esetében szokás a csoportokra csatornákként hivatkozni, melyeket a multicast címe valamint a forrás címe azonosít. A különböző végpontok ezekre a csatornákra iratkoznak fel, ezzel megkötve, hogy mely specifikus forrásoktól kívánnak multicast adatokat fogadni. A megkötött források nagyobb biztonságot jelentenek a multicast csatornáknak, miközben nagyobb követelményeket jelentenek az IGMP protokollal szemben.

Keretek felépítése

Az IGMPv3 verziója két fő típusát különbözteti meg az üzeneteknek, a tagsági lekérdezést (Membership Query), valamint a tagsági jelentést (Membership Report). A v3 lekérdezések a korábbi kettes verzióban megismerteken túl lehetnek „csoport és forrás specifikus” lekérdezések. A jelentések a korábbi funkcionalitás támogatása mellett pedig a hármas verzióban használt összetett funkcionalitással is rendelkeznek.

Az IGMPv3 keretek a többlet információk miatt komplexebbek, mint a korábbi verziók keretei. Ilyen többlet információ a források száma és címei, valamint a lekérdezés típusára vonatkozó opciók. A lekérdezések felépítése a korábban megismerten túl forrás adatokkal is rendelkezik és küldhető a 224.0.0.1-es illetve az adott csoporthoz tartozó multicast címre.

⁴ Jeff Doyle: Routing TCP/IP, Volume II, Cisco Press, 2001, 432-434 p.

A jelentések egyszerre szolgálnak csatlakozás és kilépés jelentésére is, akár egyetlen üzenetben is, köszönhetően annak, hogy a keretek a korábban már megismert attribútumokon kívül rendelkeznek úgynevezett csoport rekordokkal is, melyek az adott csoportot és forrást is azonosítják. A csoport rekordok is különböző típusúak lehetnek.⁵

- **Current-State:** Adott végpont aktuálisan csatlakozott csoportjairól és azokhoz tartozó források jelentésére. Lehetőség van úgynevezett INCLUDE és EXCLUDE csatlakozási módokra. Include mód esetében azokat a forrásokat tartalmazza a keret, amelyekről a végpont kapni szeretne adatfolyamokat, exclude mód esetén pedig azokat, amelyekről nem szeretne kapni adatfolyamokat.
- **Filter-Mode-Change:** Az INCLUDE és EXCLUDE módok közötti közvetlen váltásra szolgáló rekordok.
- **Source-List-Change:** Új források hozzáadására illetve régi források törlésére szolgáló rekordok.

Maguk a keretek a 224.0.0.22-es IGMPv3 multicast címre továbbítódnak.

Működését tekintve a hármas verzió eljárásai megegyeznek a kettes verzióéval, különbséget csak az új keretek használatából fellépő eltérések jelentenek.⁶

Csoportkezelés campus hálózatokban

A multicast csomagok a hálózaton routerről-routerre haladnak, mígnem elérnek a végpont előtti utolsó, úgynevezett last-hop routerre. Innen a router a csomagokat kiküldi az adott LAN szegmensre, amelyen a nyelők találhatóak. A végpontok valamint a router között azonban legtöbbször elhelyezkedik egy vagy több adatkapcsolati rétegbeli eszköz (például kapcsoló, amikre a továbbiakban switch-ként hivatkozok), mely nem képes betekinteni a Layer-3 csomagokba, így nem tudja, hogy a multicast adatra igényt tartó végpont melyik portján helyezkedik el. Ezek az eszközök ilyen esetben a beérkezési porton kívül minden portjukon kiküldik a beérkező multicast csomagot, amit a multicast adatra igényt nem tartó végpontok végül úgyis csak el fognak dobni. Ez a jelenség nagy mennyiségű fölösleges forgalmat generál az adott hálózatban, ami újabb szállítási rétegbeli eszközök becsatolásával tovább növekszik.

A jelenség kivédésére több lehetőség is rendelkezésre áll. Az egyik, hogy kézzel beállítjuk a multicast portokat a switch-en, ennek azonban hátránya, hogy megnöveli az üzemeltetésre fordított időt, valamint nem alkalmazkodik a multicast hálózat dinamikus változásaihoz.

A másik megoldás valamilyen dinamikus eljárás alkalmazása, melyek közül a két legjellemzőbb a Cisco hivatalos protokollja a Cisco Group Management Protocol használata illetve az IGMP Snooping.

Cisco Group Management Protocol

A CGMP használata esetén a switch-ek multicast továbbításra való felkészítése, külön CGMP üzenetek küldésével valósul meg, ahol a router kizárólag küldője, a switch kizárólag fogadója az üzeneteknek. Amikor a routerben valamilyen IGMP-vel kapcsolatos változás történik, annak eredményéről a router CGMP üzenetekkel értesíti a switch-et, aki az üzeneteknek megfelelően módosítja a CAM tábláját. A CAM tábla az egyes portokon található eszközök MAC címét tartalmazza.

⁵ Nikolay Manolov, Adel Gamil, Stephan Wong: An Investigation into Multicasting

⁶ B. Cain, S. Deering, I. Kouvelas, B. Fenner, A. Thyagarajan: Internet Group Management Protocol, Version 3 (RFC 3376)

A CGMP funkcionalitás szempontjából két üzenettípust különböztet meg, ezek a CGMP JOIN és a CGMP LEAVE, melyek felépítésükben megegyeznek. Az üzenetek a különböző multicast csoportokról és azok tagjairól tartalmaznak információt, úgymint a Group Destination Address (GDA), ami csoport MAC címét és a Unicast Source Address (USA), ami a fogadó végpont MAC címét tartalmazza.

A router CGMP JOIN üzenetek küldésével képes jelezni a switch-ek felé multicast router szerepét, valamint a végpontok csoportokhoz csatlakozását. Amikor a routerhez IGMP LEAVE üzenet érkezik, akkor a megfelelő GDA és USA értékekkel rendelkező CGMP LEAVE üzenet küldésével képes az adott nyelőhöz, csoporthoz, vagy ha szükséges, az összes csoporthoz tartozó információt kitörölni a switch CAM táblájából.⁷

IGMP Snooping

Az IGMP Snooping segítségével a switch-ek képesek betekinteni az Layer-3 IGMP üzenetekbe mielőtt azokat továbbítanák a router és a végpontok között. Ahhoz, hogy ez ne járjon túlzott erőforrás felhasználással a switch-nek hardveresen kell támogatnia az IGMP üzenetekbe történő betekintést. IGMP Snooping alkalmazásakor a switch proxyként viselkedik a végpontok és a router között.

A router által küldött Membership Query üzeneteket a switch továbbítja az összes portján, az üzenetre érkező válaszok alapján pedig, csoportonként egy választ küld a routernek. Amennyiben egy végpont csatlakozik egy csoporthoz, a switch megkapja a Membership Report üzenetet és a benne található adatok alapján módosítja a CAM tábláját. Ezek után kizárólag akkor továbbítja az üzenetet a routernek, ha még nem jelezte, hogy van csoporttag valamely portján. Amikor egy végpont kilép a csoportból, a switch a beérkező IGMP Leave portján ellenőrzi egy csoport specifikus lekérdezéssel, hogy maradt-e a porton csoporttag. Abban az esetben, ha nem maradt csoporttag, módosítja CAM tábláját, továbbá ha nem maradt több tagja a csoportnak egyik portján sem, akkor továbbítja a kilépési üzenetet a router felé.⁸

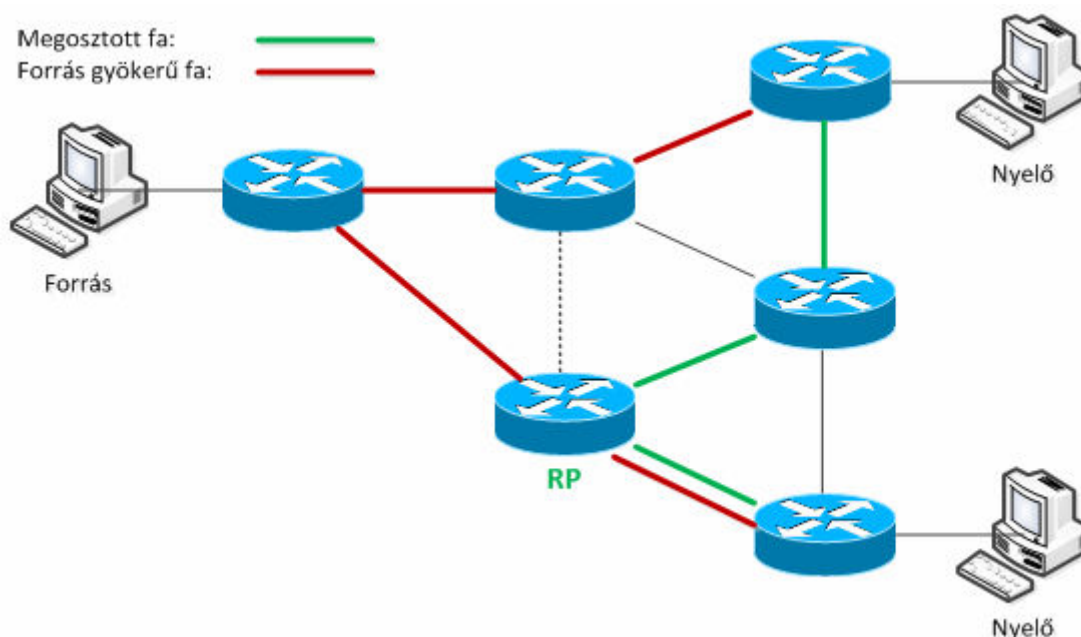
Útválasztás

A multicast továbbítás második alappillére az útvonalválasztás, melynek célja a forrás és nyelők közti optimális útvonal meghatározása. A forrás és a nyelők közötti optimális útvonal egy fa topológiaként írható le, melynek gyökere lehet a forrás vagy egy kiemelt csomópont, levelei a nyelők, csomópontjai pedig a közöttük elhelyezkedő routerek. Két szomszédos router között különbséget teszünk annak a gyökérhez viszonyított relatív távolsága alapján. A gyökérhez közelebb lévő router az ún. upstream, míg a távolabb lévő az ún. downstream router. Ez a két jelző interfészekre is hasonló módon használható.

A gyökér alapján kétféle fát különböztetünk meg: megosztott fát (shared tree) és forrás gyökerű fát (source-based tree). A source-based tree gyökere a forrás, így minden egyes forráshoz létezik egy külön fa, mely a forrás és a nyelők közötti legrövidebb utat írja le. Ez az úgynevezett Shortest Path Tree vagy SPT. Erre a fára a routerek (S,G) alakban hivatkoznak, ilyen formában jegyzik le azokat a routing táblájukba. A shared tree gyökere az adott csoporthoz tartozó (S,G) topológiák valamely közös pontja, az úgynevezett Rendezvous Point (RP). Ebben az esetben minden forrás adatfolyama ezt a gyökér és nyelők közötti fát használja. A routerek erre a topológiára (*,G) alakban hivatkoznak.

⁷ „Multicast in a Campus Network: CGMP and IGMP Snooping” Cisco dokumentáció alapján

⁸ Beau Williamson: Developing IP Multicast Networks, Volume I Cisco Press, 1999, 417-429 p.



4. ábra - Shared Tree és Source-Based Tree

Az útválasztás feladata mind unicast, mind pedig multicast továbbítás esetén a legrövidebb út meghatározása. Míg unicast esetében ez a nyelőhöz vezető legrövidebb út, addig multicast esetén, az adattovábbításban egyetlen konstans pont, a forráshoz vagy RP-hoz vezető legrövidebb út. E miatt a fordítottság miatt is nevezik a multicast továbbítást „reverse path forwarding”-nak.

Útválasztás során a protokollok feladata, hogy meghatározzák az adott routeren, a forráshoz legközelebb eső (upstream) interfészt, valamint a nyelőkhöz vezető (downstream) interfészeket, majd az ezekből képezett információkat elraktározzák a routing táblájukban. Mivel az adatfolyam fennállása során nyelők léphetnek ki illetve be a csoportba, ezért a routing táblák alapján meghatározott topológiák dinamikusan változhatnak.

Kétféle topológiát különböztetünk meg, sűrű (dense) illetve ritka (sparse) topológiát. Sűrű topológiát a nyelők nagy százaléka jellemzi az összes végponthoz képest, míg ritka topológiát az ellenkezője. A sűrű topológiák többnyire flood-and-prune módszer alapján működnek, ahol a downstream interfészek közé az összes nem upstream interfész bekerül, azaz implicit módon csatlakoznak a downstream routerek a fába. Ez a hálózat egészének elárasztásához vezet. Ahhoz, hogy a multicast adatfolyamok ne haladjanak olyan irányba, ahol nem találhatóak nyelők, az ott található routerek erről, Prune eljárás alkalmazásával értesítik az upstream routereiket. A „prune”-olt (csonkolt) downstream interfészekben a routerek nem továbbítják a multicast forgalmat, egészen addig míg a prune időzítő le nem jár. Amikor az időzítő lejár az egész prune eljárás újraindul. Ez az implicit csatlakozás és a prune eljárás ismétlődése többlet erőforrás-felhasználást eredményez.

A ritka topológia esetén nem beszélünk elárasztásról, a downstream routerek explicit módon csatlakoznak a fához, a hozzájuk beérkező IGMP üzenetek alapján.⁹

Distance Vector Multicast Routing Protocol

A DVMRP, mint ahogy a neve is mutatja egy távolság vektoron alapuló útvonalválasztó protokoll, mely a RIP unicast útválasztó protokoll egy változatát használja a forráshoz vezető

⁹ Jeff Doyle: Routing TCP/IP, Volume II, Cisco Press, 2001 450-460 p.

legrövidebb út meghatározásához. Az így nyert információk alapján kialakított topológiát figyelembe véve a DVMRP egy sűrű módú protokoll, mely flood-and-prune alapon működik és minden egyes forráshoz külön forrás-gyökerű fát hoz létre. A DVMRP működéséhez hét különböző üzenetet használ.

Szomszédsági kapcsolatok

Első lépésként a routereknek szükségük van minden szomszédos routerük megismerésére. Ennek teljesítésére szolgál a DVMRP Probe üzenet, melyet minden router elküld aktiválásakor, valamint azt követően megadott időközönként, a 224.0.0.4 „minden DVMRP router”-t címző multicast címre. A DVMRP Probe üzenet tartalmazza egy a router DVMRP képességeit leíró 8 bites értéket, egy Generation ID-nak nevezett 32 bites attribútumot, valamint a küldő router kimeneti interfészén található, ismert szomszédos routerek listáját. A probe üzenetben található Generation ID változása hirdeti az upstream router felé, hogy a router újraindult és a tőle korábban kapott prune üzenetek semmisnek tekinthetőek. Két szomszédos router egy-egy probe üzenet váltásával egy kétirányú kapcsolatot alakít ki.

Topológia kialakítása

A routerek a DVMRP routing táblában található információkat, megadott időközönként DVMRP Report üzenetekben továbbítják a 224.0.0.4-es címre. A report üzenet tartalmazza a router által ismert forráshálózatok címét valamint az ezekhez tartozó hálózati maszkokat és metrikákat. Mivel a DVMRP egy távolságvektoros protokoll a metrikák ugrás számok. Az 1 és 31 érvényes metrikák, melyek az interfészen található upstream routert jelölnék, míg a 33 és 63 közötti metrikák úgynevezett útvonalfüggőségek, melyek downstream routert jelölnék. A 32 a végtelen metrika, ami többek között a multicast topológia maximális kiterjedését is meghatározza. A routerek az upstream router felé a forráshoz tartozó és a végtelen metrika összegeként létrejövő távolsággal jelzik, hogy downstream routerek. A multicast hálózat konvergenciáját követve minden router routing táblájában megtalálható, hogy mely interfészek bejövő és melyek kimenő interfészek egy adott forrás által küldött adatfolyam továbbításakor.¹⁰

Adatfolyam továbbítása

Multicast adatfolyam továbbításakor az adatokat fogadó routerek egy úgynevezett RPF ellenőrzést végeznek. Az RPF ellenőrzés lényege, hogy megállapítsák, hogy a multicast adatok a megfelelő (a routing táblában a forráshoz tartozó upstream) interfészen érkeznek-e be. A nem megfelelő interfészen beérkező csomagokat eldobják, hogy ne alakulhasson ki hurok az adattovábbításban. Amennyiben a megfelelő upstream interfészen érkeznek az adatok, akkor a továbbítási táblában feljegyzésre kerül az (S,G) fa és a downstream interfészekon megtörténik a továbbítás. Abban az esetben, ha a routernek nincsenek downstream routerei, valamint IGMP üzenetekkel megállapítja, hogy nincsenek közvetlen nyelői a multicast csoportnak egyetlen interfészén sem, akkor DVMRP Prune üzenetet küld az upstream szomszédjának. A DVMRP Prune üzenetek tartalmaznak egy időértéket (alapbeállítások szerint 2 óra) ameddig az upstream router prune állapotban tartja az adott interfészen az (S,G) által meghatározott fát és ameddig nem továbbítja a hozzá tartozó multicast adatfolyamot azon az interfészen. Amennyiben új tagok csatlakoznak a csoporthoz a kapcsolódó router DVMRP Graft üzenetet küld az upstream szomszédjának. A graft üzenet ugrásonként addig továbbítódik, amíg egy aktívan továbbító routerig el nem ér. A graft üzeneteket a routerek DVMRP Graft Ack üzenettel nyugtázzák annak érdekében, hogy megkülönböztethető legyen a graft üzenet elveszése a forrástól érkezett adatfolyam hiányától.

¹⁰ D. Waitzman, C. Partridge, S. Deering: Distance Vector Multicast Routing Protocol (RFC 1075)

A maradék két üzenettípus a DVMRP Ask Neighbors2 és az arra érkező DVMRP Neighbors2 válaszüzenet melyek hibaelhárításnál használatosak.¹¹

Értékelés

A DVMRP az első multicast routing protokoll, melyet leginkább az MBONE internetes multicast maghálózatban használnak. A protokoll hátránya, hogy használata során csak korlátozott mértékig növelhető a multicast topológia kiterjedése, valamint, hogy használatakor szükséges egy külön multicast routing tábla karbantartása. Emellett még a sűrű módot többlet erőforrás-felhasználás is jellemzi.

Multicast OSPF

Open Shortest Path First

A MOSPF, a unicast Open Shortest Path First útválasztó protokoll multicast kiegészítése. Az OSPF egy link-state protokoll, ahol a routerek különböző típusú „link state advertisements” (LSA) üzenetekkel értesítik (*Hello* üzenetekkel felfedezett) szomszédjaikat interfészeikről és azok állapotáról. A szomszédos routerek ezeket az információkat elraktározzák a link-state adatbázisukban, majd továbbküldik az LSA-t a többi interfészükön. Mikor a routerek link-state adatbázisai szinkronizálódtak, a routerek az SPF algoritmussal kiszámítják az összes lehetséges célhoz, a legrövidebb utat meghatározó, hurokmentes fát (Shortest Path Tree: SPT). A továbbiakban a routerek *Hello* üzenetekkel tartják fent szomszédjai kapcsolataikat, és megadott időközönként elküldik LSA-ikat. Az SPF algoritmussal kiszámolt topológia kiterjedése az OSPF terület (OSPF area) határáig ér. Az OSPF terület az az adminisztratív terület, amelyen belül az OSPF üzenetek küldése és az SPT kiszámítása megtörténik.

Multicast kiegészítések

Ahhoz, hogy a MOSPF alkalmas legyen multicast forgalom továbbítására, az OSPF-et egy új, Group Membership LSA-val (LSA type 6), a Hello üzenetek *Options* mezőjét egy MC multicast képességeket jelző bittel, a Router LSA *rtype* mezőjét pedig egy W bittel egészítették ki, aminek bebillentésével a router jelzi, hogy az összes multicast forgalomra igényt tart.¹²

Működés a MOSPF területen belül

A MOSPF egy sűrű topológiájú protokoll, melynek előnye a DVMRP-hez képest, hogy explicit csatlakozást valósít meg a routerek fába történő bekapcsolódásakor. Amikor egy végpont IGMP üzenettel jelzi, hogy csatlakozott egy csoporthoz, akkor az adott szegmensen a továbbításért felelős ún. Designated Router létrehoz egy bejegyzést a helyi csoport adatbázisába a csoport címével és a subnet címével ahonnan az IGMP-t kapta. Ezek után csoportonként egy Group Membership LSA elterjesztésével a MOSPF területben, explicit módon jelzi, hogy vannak a csoportokhoz tartozó nyelők az interfészein. Amikor a routerek adatbázisai szinkronizáltak, azok készen állnak a továbbítási fa kiszámítására, azonban ez csak az első multicast csomag érkezésekor történik meg, ugyanis csak ekkor szereznek tudomást a fa gyökeréről szolgáló forrásról. A tényleges multicast adatfolyam továbbítása során, mivel az SPF algoritmus biztosan hurokmentes fát számít ki, nincs szükség RPF ellenőrzésre.

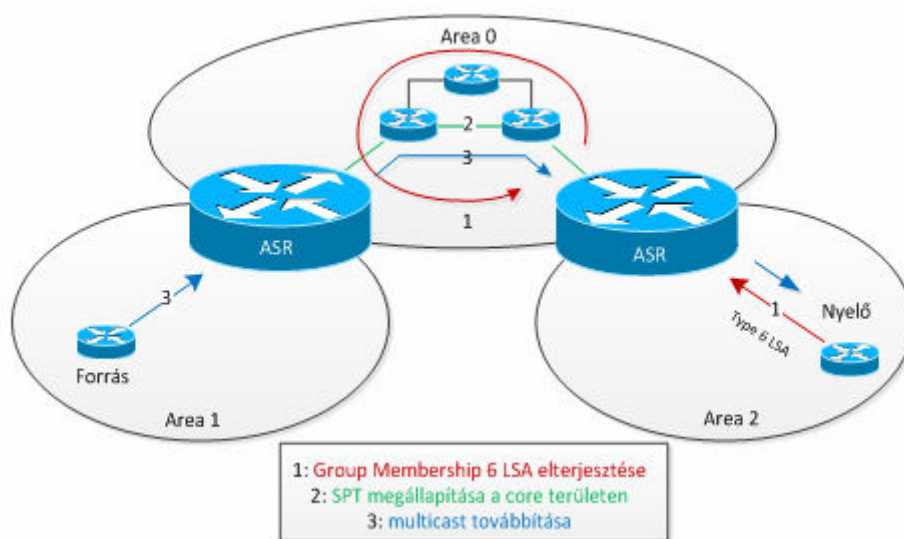
Működés MOSPF területek között

A fenti eljárás kizárólag az OSPF területen belül érvényesül. Ahhoz, hogy a multicast

¹¹ Jeff Doyle: Routing TCP/IP, Volume II, Cisco Press, 2001 464-473 p.

¹² J. Moy: MOSPF: Analysis and Experience (RFC 1585)

területek között is működjön szükséges egy MOSPF Area Border Router (ABR), ami a gerinc (backbone) terület és valamely nem gerinc terület között helyezkedik el, és amely az területek közötti multicast továbbítóként működik. Az ABR a beérkező type 6 LSA-k alapján értesül a nem gerinc területben lévő csoporttagságokról, majd ezek alapján Group Membership LSA-t küld a gerinc felé, így az összes a gerincben lévő router tudni fog a nem gerinc területek csoportjairól. A gerincben minden csoporthoz kiszámításra kerül egy SPT. Mivel ABR-ek nem küldenek Group Membership LSA-t a nem gerinc hálózatokba, így azok nem tudnak a területen kívüli csoporttagokról. Ahhoz, hogy a tényleges multicast forgalom kiküldésre kerüljön a gerinc területre, az ABR-nek be kell billentenie a W wildcard bitet, hogy minden multicast forgalmat megkapjon, majd azt tovább tudja küldeni a gerinchálózaton számított SPT-k alapján.¹³



5. ábra - Területek közötti MOSPF működés

Értékelés

A MOSPF az OSPF alapok miatt egy jóval kifinomultabb multicast útválasztó megoldás, mint a DVMRP. A multicast hálózat kiterjedését tekintve a MOSPF jóval nagyobb kiterjedés elérésére alkalmas, aminek a karbantartását megkönnyítheti a különböző területek alkalmazása. Működését tekintve a MOSPF továbbra is sűrű topológiában működik megfelelően, ugyanis a topológia dinamikus változása az SPF algoritmus újbóli meghívását igényli minden egyes változásnál, ami jelentős erőforrás felhasználást von maga után.

Core Based Tree

A CBT egy protokoll-független multicast útválasztó eljárás, mely nincs semmilyen meghatározott unicast protokollhoz kötve. A CBT mint a neve is sejteti egy megosztott fa topológiát kiépítő protokoll, amely ritka módban működik. Protokoll függetlenségének köszönhetően szabadon használható bármely unicast útválasztó protokollal.

Topológia kialakítása

Működését a CBT kilenc különböző üzenettel valósítja meg, melyek a 224.0.0.15-ös címre továbbítódnak. A CBT a topológia kialakítása során a routerek explicit módon csatlakoznak a multicast fához, melynek gyökere a Core vagy mag. A megosztott fa gyökerét képező mag

¹³ Beau Williamson: Developing IP Multicast Networks, Volume I, Cisco Press, 1999, 177-187 p.

ismerete szükséges a topológia kialakításához. A gyökér router címét kézzel be lehet konfigurálni a hálózat routereiben, így az a unicast útválasztási táblából megtalálható. A mag változása azonban ilyenkor magas karbantartási követelményekkel jár. Ennek automatizálására használható a Bootstrap algoritmus, amelynek során több mag jelölt a közöttük váltott üzenetekkel eldönti, hogy ki lesz a mag és ezt elterjesztik a hálózat routerei között.

Amikor egy router aktívvá válik a multicast hálózatban, egy *Hello* üzenetet küld ki minden multicast képes interfészén, mely tartalmazza az interfészhez tartozó preferencia értékét, amely alapján az adott szegmensen lévő routerek meghatározzák, hogy ki lesz a Designated Router. A kisebb érték magasabb preferenciát jelent, így a legkisebb értékű router lesz a DR. A DR feladata, hogy adott időközönként, valamint újonnan megjelenő routerek *Hello* üzeneteire válaszol, *Hello* üzenetet küldjön.

Amikor egy nyelő csatlakozik egy csoporthoz és erről IGMP üzenetben értesíti a DR-t, az a unicast útválasztási táblájából megkeresi, hogy a mag felé melyik interfészén kell továbbítani a csomagokat, majd elküld egy JOIN_REQUEST üzenetet a mag felé. Ez az üzenet tartalmazza a mag címét, a csoport címét valamint a küldő router címét. A szomszédos router a beérkező üzenet alapján eldönti, hogy benne van-e a fában vagy sem, illetve, hogy mag router-e. Amennyiben benne van a fában úgy egy JOIN_ACK üzenettel nyugtázza a kérelmet és megkezd a multicast adatok továbbítását, amennyiben nem, úgy ő is egy JOIN_REQUEST üzenettel próbál meg csatlakozni a fához.

Topológia karbantartása

Amennyiben a fa kiépült, a downstream router (child) ECHO_REQUEST üzenetekkel tartja fent a kapcsolatot az upstream routerrel (parent). Ezekre az üzenetekre az upstream router ECHO_REPLY üzenetekkel válaszol, mely tartalmazza az adott interfészen az összes általa továbbított csoport címét. Amennyiben ez az üzenet nem érkezik meg, vagy nem tartalmaz olyan csoportokat, amelyeket a downstream router kiszolgál, a downstream router QUIT_NOTIFICATION üzenetet küld az upstream router felé, ezzel leszakítva magát a fáról. Ezzel egy időben az érvénytelen csoportok listáját tartalmazó FLUSH_TREE üzenetet küldd a downstream interfészein, melyek a fa alsóbb részeit elárasztva, lebontják azokat. Amennyiben egy last-hop routerről eltűnik az utolsó csoporttag is, szintén a QUIT_NOTIFICATION üzenet küldésével válhat le a fáról a router.¹⁴

Adatfolyam továbbítása

Adattovábbítás során a CBT különbséget tesz azon két eset között, amikor a megosztott fa tagja a forrás, illetve a megosztott fán kívüli a forrás. Mivel a CBT fa kétirányú így a fa levelétől a gyökeréig is történhet adattovábbítás, így az első eset nem okoz problémát a CBT számára. Amikor a forrás nem tagja a megosztott fának, akkor a hozzá kapcsolódó last-hop router egy IP-IP beágyazással egy alagutat hoz létre a maghoz, amelyben unicast módon továbbítja a csomagokat.

Értékelés

A CBT ritka módja, valamint a megosztott fája miatt előnyösebb multicast hálózat kialakítására, mint a korábbi két topológia. A hálózat kiterjedését nem korlátozza csak a használt unicast útválasztó protokoll. Erőforrások felhasználása szempontjából is előnyös a CBT, mert nem használ a unicaston túl további útválasztási protokollokat. Hátránya, hogy sok, nagy forgalmú csoport esetén az összes forgalom a mag körül csoportosul ezért mind a

¹⁴ Tony Ballardie, Paul Francis, Jon Crowcroft: Core Based Trees (CBT) - An Architecture for Scalable Inter-Domain Multicast Routing

mag router, mind pedig a környező kapcsolatok szűk keresztmetszetet jelenthetnek az adattovábbításban.

Protocol Independent Multicast

A PIM protokoll, mint ahogy a neve is mutatja egy a unicast útválasztási protokolltól független multicast protokoll. A protokollnak két külön változata van, a PIM-DM mely sűrű topológiákhoz illeszkedik megfelelően, illetve a fejlettebb PIM-SM, ami a ritka topológiákhoz. A protokollnak két verziója van PIMv1 és PIMv2. Míg az első verzió IGMP csomagokba ágyazva továbbítja üzeneteit a 224.0.0.2-es multicast címre, addig a második verzió saját IP-be ágyazott csomagokat használ, amiket a 224.0.0.13-as címre továbbít.

Sűrű mód

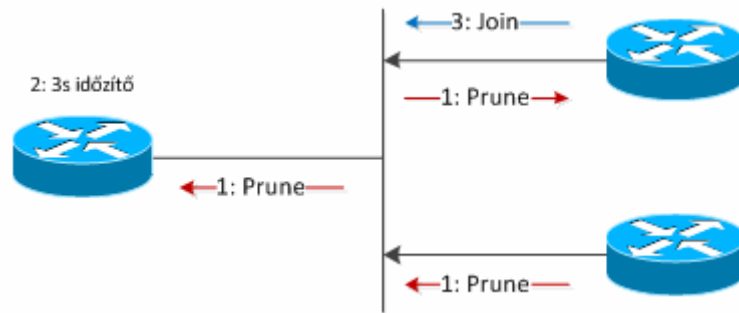
A PIM-DM egy sűrű topológiákhoz alkalmazkodó protokoll, mely implicit csatlakozást feltételez, ezért működésében leginkább a DVMRP-re hasonlít. A topológia kiépítésére és karbantartására a PIM-DM öt különböző üzenetet használ.

- Hello üzenet (PIMv1 esetén Query): a szomszédsági kapcsolatok kiépítésére használatos
- Join/Prune üzenet: fára történő csatlakozásra és a fáról történő leválásra használatos
- Graft és Graft-Ack üzenet: a fához történő újracsatlakozásra valamint ennek a nyugtázására használatos
- Assert üzenet: a továbbító router kiválasztására szolgál megosztott szegmenseken

A topológia kialakítása

A topológia kialakításának első lépése itt is, mint a DVMRP-nél a szomszédsági kapcsolatok felfedezése. Amikor a hálózatban egy PIM-re konfigurált router aktívvá válik (majd ezt követően 30 másodpercenként), Hello üzenetet küld minden interfészén. A Hello üzenet egy holdtime értéket tartalmaz, ami megmondja a szomszédos routernek, hogy mennyi az az idő, ami alatt, ha nem kap újabb *Hello* üzenetet a küldő routerrel a szomszédságot semmisnek tekintheti. Ez a holdtime idő a *Hello* üzenetek között eltelt idő 3,5 szerese. A szomszédsági kapcsolatokon kívül a hello üzenetek segítségével történik meg a Designated Router kiválasztása a többszörös hozzáférésű szegmenseken, az IGMPv1 támogatása végett.

Mivel a PIM-DM elárasztásos működést valósít meg, abban a pillanatban, amikor egy multicast csoportnak egy forrás adatot küld, az összes végpontot elérő fa topológia jön létre. Amikor valamely router megkap egy multicast csomagot, a multicast továbbítási táblájában létrehoz egy (S,G) elemet (mely tartalmazza az bejövő és kimenő interfészeket), majd kiküldi a csomagot az összes kimeneti interfészén. Ezzel az elárasztással a forgalom eléri az összes last-hop routert. Amennyiben egy router nem tart igényt, a forgalomra egy Join/Prune üzenetet küld, mely tartalmazza a csoportok címeit, amiket az upstream routernek prune állapotra kell állítania a küldő router felőli interfészén. A Prune állapot, mint a DVMRP-nél, itt is timeout-ol, ezért frissíteni kell. Abban az esetben, ha az upstream routernek a downstream interfészén többszörös hozzáférésű szegmens van és több router is található rajta, akkor Prune üzenet beérkezésekor a forgalomra továbbra is igényt tartó routereknek aktívan jelezniük kell egy Join/Prune üzenettel ezt az upstream router felé. Az upstream router egy 3 másodperces időkorlátot tart fent ennek beérkezésére, amelynek lejártakor állítja csak az interfészén a csoportot prune állapotba.



6. ábra - Prune eljárás osztott szegmensen

Amennyiben a routernek mégis szüksége lenne a multicast csoport forgalmára, egy Graft üzenettel jelezheti azt a szomszédos routernek, amire az a Prune állapot törlésével és egy Graft-Ack üzenettel reagál. Mind a Graft mind a Graft-Ack üzenet unicast módon továbbítódik a routerek között.

Abban az esetben ha a unicast topológia megváltozik a unicast routing tábla tükrözni fogja a változást és a PIM értesülni fog a változásról. A routerek ilyenkor a megfelelő interfészekén keresztül küldött prune és graft üzenetekkel alkalmazkodnak a topológiai változásokhoz.

Adatfolyam továbbítása

A multicast csomagok továbbítása routerről-routerre történik a multicast továbbítási táblázatokban található (S,G) bejegyzések alapján. A PIM RPF ellenőrzést végez a beérkező csomagokra, és ha azok a megfelelő upstream interfészen érkeztek be, továbbítja azokat. Az egyetlen nehézséget, a többszörös hozzáférésű szegmenseken tapasztalhatjuk, ahol mindegyik router továbbítja a multicast csomagokat, így azok több példányban jelennek meg a szegmensen. Ennek kivédésére szolgál az úgynevezett kiválasztott továbbító (designated forwarder) kiválasztása. Amennyiben egy router az adott csoporthoz tartozó forgalmat valamely kimeneti interfészen kapja meg, akkor egy Assert üzenetet küld ki az interfészen. Az Assert üzenet tartalmazza a csoport és a forrás címét valamint a forráshoz vezető út költségét és adminisztratív távolságát. A szegmensen található routerek mindegyike küld egy ilyen Assert üzenetet és elsődlegesen az adminisztratív távolság majd a költség alapján meghatározásra kerül a továbbító személye. Abban az esetben, ha ez az értékek egyenlősége miatt nem vezet megoldásra, a legmagasabb IP címmel rendelkező router lesz a továbbító.¹⁵

Értékelés

A protokoll előnye, hogy bármely unicast útválasztó protokollal használható így csak azok szabnak határt a hálózat kiterjedésének. A sűrű topológiájú multicast hálózatokban jól használható, azonban rendelkezik az ezzel együtt járó hátrányokkal és erőforrás szükségletekkel is.

Ritka mód

A ritka módú topológiáknál használatos a PIM-SM, mely azonban sűrű topológiákat is jól kezeli. Nagy előnye, hogy explicit csatlakozást alkalmaz és mind megosztott mind pedig forrás-gyökerű fát is alkalmaz. A forrás-gyökerű fának köszönhetően nem szükséges minden multicast csomagnak a PIM-SM-nél a megosztott fa, randevú pontnak (RP) nevezett gyökeréig továbbítódnia, ha a nyelő egy rövidebb útvonalon is elérhető.

A PIM-SM hét üzenettípust használ, melyek közül három a PIM-DM-ben is használt Hello, Join/Prune és Assert üzenet. A további üzenetek:

¹⁵ Jeff Doyle: Routing TCP/IP, Volume II, Cisco Press, 2001, 496-509 p.

- Bootstrap üzenet: bootstrap eljárás vezérlőüzenetei
- Candidate-RP-Advertisement üzenet: randevú pont hirdetésére szolgáló üzenet
- Register és Register-Stop üzenet: forrás által küldött csomagok kezelésére szolgál

A topológia kialakítása

A PIM-SM esetében is mint a PIM-DM-nél az első lépés a topológia kialakításában a szomszédsági kapcsolatok feltárása. Ezeknek a kapcsolatoknak a feltárása valamint a többszörös hozzáférésű hálózatokon a designated router és forwarder kiválasztása a PIM-DM-nél megismertek szerint történik.

A fa felépítése a randevú pont meghatározásával történik. Minden csoporthoz tartozik egy randevú pont, melynek címét a hálózatban lévő PIM-SM routereken konfigurálhatjuk, vagy valamely automatikus eljárással terjeszthetjük el. Ezek az eljárások a CBT-nél már említett Bootstrap valamint a Cisco Auto-RP eljárása.

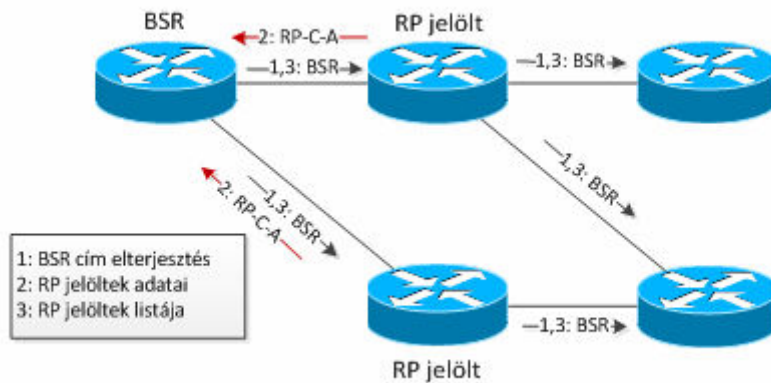
Amikor egy multicast csoport nyelője IGMP üzenettel jelzi egy csoporthoz csatlakozását, az azt fogadó router létrehoz egy (*,G) bejegyzést a multicast továbbítási táblájában, ahol feljegyzi az IGMP üzenet beérkezési interfészét, mint kimeneti interfész. Ezután a routeren tárolt RP információk alapján, a unicast útválasztási táblából megkeresi a randevú ponthoz vezető interfészt, melyet mint a bejövő (RPF) interfész tárol el. Az (*,G) elem rögzítése után a router Join/Prune üzenetet küld az upstream router felé. A Join/Prune üzenet tartalmazza a csoport és az RP címét is, így a fa szintről-szintre felépülhet egészen az RP-ig, vagy a már meglévő megosztott fa legközelebbi pontjáig. Az upstream routerek abban az esetben, ha nem tartalmazzák a (*,G) elemet, vagyis nem tagjai a megosztott fának, létrehozzák azt és tovább csatlakoznak. Abban az esetben, ha a csoportnak még nem volt egyetlen vevője sem a csatlakozás egészen a randevú pontig folytatódik. Az RP router létrehozza a (*,G) bejegyzést, melyben kimeneti interfészt az előzőek szerint állítja be, a bemeneti interfészt Null értékre állítja, az RPF szomszéd értéket pedig 0.0.0.0-ra állítja, ezzel jelezve, hogy ő az RP. A CBT-vel ellentétben a PIM-SM-nél nincs szükség a forrástól érkezett multicast csomagokra, hogy a megosztott fa kiépüljön, úgyhogy a topológia ezzel készen is áll. Ha az adott ágon a nyelők megszűnnek, Join/Prune üzenetekkel történik a leválasztás.¹⁶

RP meghatározása

Bootstrap eljárás esetén úgynevezett RP jelölteket határozunk meg a hálózaton melyek önmagukat címükkel és RP prioritásukkal hirdetik a Bootstrap Router (BSR) felé, unicast-tal küldött Candidate-RP-Advertisement üzeneteken keresztül. Ahhoz, hogy ez működjön értesülniük kell a BSR routerről. Az RP jelölteken kívül BSR jelölteket is kijelölünk a hálózaton, melyek prioritásukat és címüket tartalmazó Bootstrap üzenetekkel hirdetik magukat a hálózaton. A Bootstrap üzenetek hopról-hopra elárasztják a hálózatot így minden PIM képes router megkapja azokat. A Bootstrap jelöltek közül a legmagasabb prioritású vagy másodlagosan a legmagasabb című lesz a BSR. A BSR feladata, hogy az RP jelöltektől kapott információk alapján előállítson egy listát a csoportokhoz tartozó RP jelöltekből, majd azt Bootstrap üzenetek segítségével megadott időközönként elterjessze a hálózaton. A hálózat routerei ebből a listából határozzák meg a csatlakozott csoport RP-jét, a kisebb RP prioritás, másodlagosan egy számított hash érték, harmadlagosan pedig a magasabb IP cím alapján.¹⁷

¹⁶ B. Fenner, M. Handley, H. Holbrook, I. Kouvelas: Protocol Independent Multicast - Sparse Mode (PIM-SM) Protocol Specification (Revised) (RFC 4601)

¹⁷ N. Bhaskar, A. Gall, J. Lingard, S. Venaas: Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM) (RFC 5059)

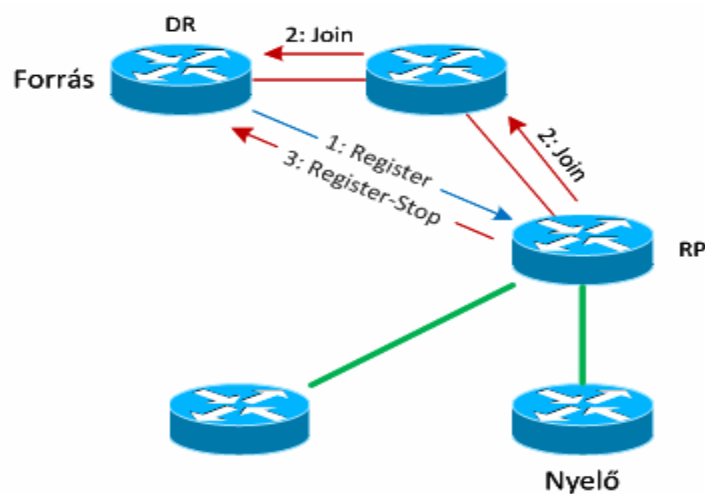


7. ábra - BSR mechanizmus

Az Auto-RP a Cisco protokollja, ami hasonlóan működik a bootstrap-hez. A hálózaton itt is ki kell jelölni RP jelölteket, amik üzeneteit a BSR-hez hasonló funkcionalitású „RP mapping agent” fogadja. Az RP jelöltek RP-Announce üzeneteiket a 224.0.1.39-es címre küldik, amin az „RP mapping agent” figyel. A hálózatban egyetlen ilyen mapping agent van, itt nincsenek külön jelöltek. A mapping agent a BSR-rel ellentétben nem egy listát küld a lehetséges randevú pontokról, hanem meghatározza a csoport-RP párokat és azokat RP-Discovery üzenetekben küldi el a 224.0.1.40-es címre, amelyre minden Cisco PIM-SM router figyel.¹⁸

Adatfolyam továbbítása

Ahogy a CBT-nél is, a PIM-SM-nél is el kell juttatni valahogy a multicast csomagokat a megosztott fa gyökeréhez, a randevú ponthoz. Mivel a PIM, RPF ellenőrzést használ, nem lehetséges a megosztott fán a kétirányú multicast áramlás, így ez nem oldható meg a forrás, fába csatolásával. Amikor a forráshoz legközelebb lévő router multicast csomagot kap, megkeresi a hozzá tartozó eltárolt RP címét, majd Register üzenetekbe ágyazva, unicast módon elküldi azt az RP felé. Az RP amennyiben van a csoportnak valahol a megosztott fában vevője, egy (S,G) forrás-gyökerű fa (SPT) kiépítését kezdeményezi a forrás felé Join/Prune üzenet küldésével. Amikor az SPT kiépült és a multicast forgalom azon továbbítható, akkor az RP egy Register-Stop unicast üzenetet küld a forrás Designated Router-hez.



8. ábra - SPT kiépítése RP és DR között

¹⁸ Beau Williamson: Developing IP Multicast Networks, Volume I, Cisco Press, 1999, 318-332 p.

Abban az esetben, ha a Register üzenet érkezésekor nincsen a megosztott fán nyelő, akkor az RP Register-Stop üzenetet küld vissza válaszul, ezzel jelezve, hogy ne küldje a DR feleslegesen felé a multicast forgalmat. A hiányzó nyelők ellenére az RP megtartja a létrehozott (*,G) bejegyzést a továbbítási táblájában Null kimeneti értékkel, hogy egy nyelő csatlakozásakor egyből kiépíthesse a forráshoz az SPT-t. Az RP hibák esetére gondolva a Register-Stop állapot timeout-ol a DR-ben ezért az, megadott időközönként új Register üzeneteket küld.

Bizonyos esetekben a megosztott fa alkalmazása többletutakhoz vezet, mint például abban az esetben, ha a forrás DR-e és a nyelő között rövidebb út is van, mint az RP-n keresztül vezető. A PIM-SM képes a megosztott fáról SPT-re átállni bizonyos esetekben. A nyelő felé eső last-hop router a beérkező multicast csomagok vételekor, megvizsgálhatja a forráshoz vezető legrövidebb utat, és SPT felállítását kezdeményezheti a forrás DR felé. Az SPT kiépülése után a last-hop router, Join/Prune üzenettel lecsatlakozik a megosztott fáról és a továbbiakban az SPT-n keresztül kapja meg a multicast adatokat.

A megosztott fáról az SPT-re való átállást egy forgalomküszöb szabályozza, melynek átlépésénél a last-hop router kezdeményezi az SPT kiépítését. Amennyiben a forgalom a meghatározott küszöb alá esik, úgy a forgalom visszaáll a megosztott fára.¹⁹

Értékelés

A PIM-SM a ritka topológia és az explicit csatlakozás miatt előnyösebb protokoll mint a korábbiak, mely előnyt a megosztott és forrás-gyökerű fák együttes alkalmazása tovább növel. Előnyös, hogy a unicast útválasztási protokolltól való függetlenség miatt a multicast topológia kiterjedésének nincs külön korlátja. A hálózat növekedését nem gátolja, hogy a hálózati eszközöknek szükséges az RP ismerete, köszönhetően annak automatikus elterjesztését végző eljárásoknak. A megismert protokollok közül a PIM-SM nevezhető a legfejlettebb és leghatékonyabb protokollnak.

¹⁹ Jeff Doyle: Routing TCP/IP, Volume II, Cisco Press, 2001, 514-530 p.