

Hálózati Technológiák és Alkalmazások

Vida Rolland, BME TMIT

2020. november 24.






Csoportos kommunikáció

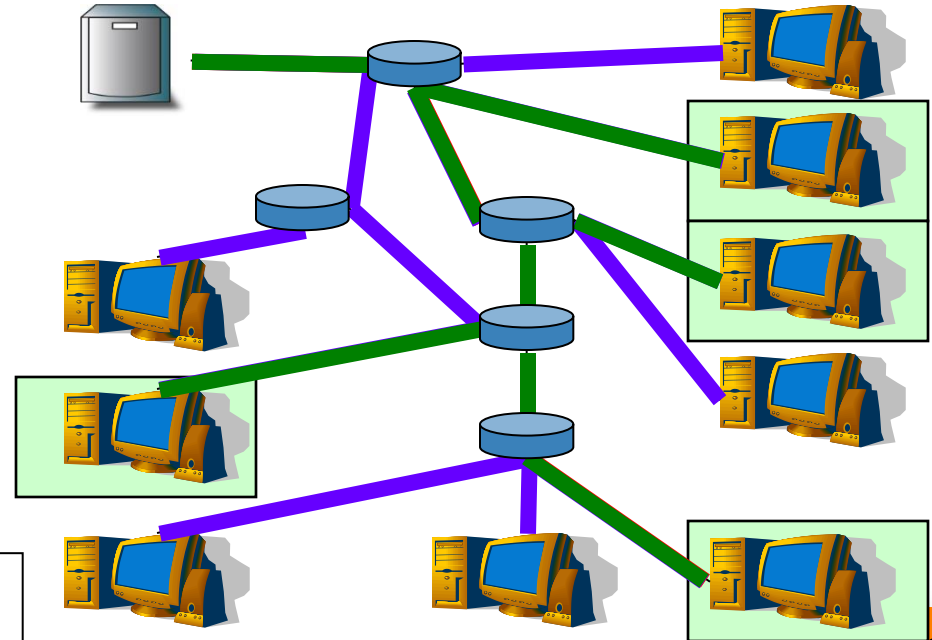
- Cél: egy egyedi célállomás helyett egy célállomás halmazzal (csoporttal) kommunikálni
 - „természetes” általánosítása a pont-pont kommunikációnak (unicast)
- **Multicast** = többesadás



Mi is a multicast?

- **Unicast**
 - Pont - pont
 - Cél cím: egyedi vevő címe
- **Broadcast**
 - Pont - mindenki (csomagszórás)
 - Cél cím: a hálózat címe
- **Multicast**
 - (Több)pont - többpont
 - Cél cím: a csoport címe

Unicast csomagküldés	
Broadcast csomagküldés	
Multicast csomagküldés	



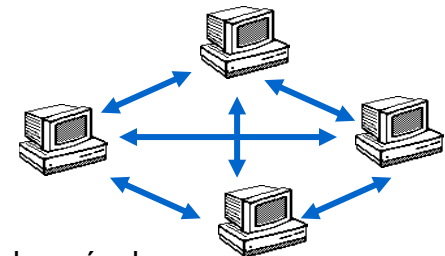
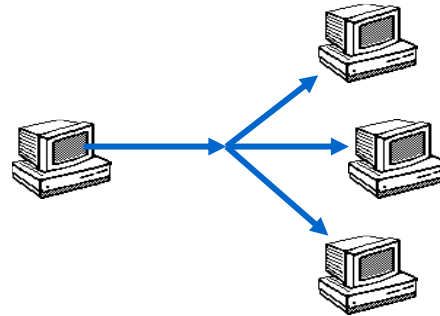
Csoportos kommunikáció

- A csomagokat egy csoport minden tagjához el kell juttatni, nem csak egy célállomáshoz
 - A csoport felépítése (tagsága) dinamikus lehet
- Működési alapelv: miután egy csoport létrejön
 - Érdeklődő vevők (receiver) csatlakoznak a csoporthoz
 - A hálózat foglalkozik a csoport karbantartásával és a csomagok továbbításával



Multicast alkalmazások

- Számos alkalmazás nem pont-pont alapú
 - Pont-többpont
 - Távoktatás
 - Cache update
 - Video on demand
 - Többpont-többpont
 - Videokonferencia, Audiokonferencia, Chat,
 - Elosztott hálózati játékok
 - Kooperatív alkalmazások



Követelmények

- Nincs olyan megoldás, mely minden környezetben ideális lenne
- A követelmények nagyban változnak
 - az alkalmazás igényeitől függően
 - a csoport méretétől függően
 - a hálózati szolgáltatásoktól függően
 - a résztvevők heterogeneitásától függően



Részvételi feltételek

- Csoporttagság ellenőrzés
 - Nyitott csoport: bárki lehet tag
 - Zárt csoport: tagsági korlátozások

- Forrásellenőrzés
 - Bárki küldhet csomagot a csoportnak
 - Csak egy csoporttag küldhet csomagokat
 - Csak egy kijelölt forrás küldhet csomagokat



Adatátviteli feltételek

- **Pont-pont kommunikáció**
 - megbízható vagy best-effort (garanciák nélkül)
 - A cél ellenőrzi a kapott csomagot: vagy OK, vagy nem
- **Pont-multipont kommunikáció**
 - minden vevő másként értékelheti a szolgáltatást
- **Különböző megbízhatósági szintek**
 - 0-megbízhatóság: egyetlen vevőnek sem garantált a megbízható átvitel
 - 1-megbízhatóság: legalább egy vevőnek garantált
 - k-megbízhatóság: legalább k vevőnek garantált
 - totális megbízhatóság: minden vevőnek garantált a megbízható szolgáltatás



Multicast különböző rétegekben

- A multicast (többesadás) szolgáltatást különböző rétegekben lehet implementálni
 - Adatkapcsolási réteg (data link layer)
 - pl. Ethernet multicast
 - Hálózati réteg (network layer)
 - pl. IP multicast, Xcast
 - Alkalmazási réteg (application layer)
 - pl. Narada, TBCP
- Melyik megoldás a jobb?
 - Attól függ, nincs általános megoldás

Multicast különböző rétegekben

Rétegek

Alkalmazási réteg

Middleware

Szállítási réteg

Hálózati réteg

Adatkapcsolás

Unicast

Alkalmazások

Overlay Multicast

TCP

IP

Ethernet Unicast/Multicast

Multicast

Reliable Multicast

IP Multicast

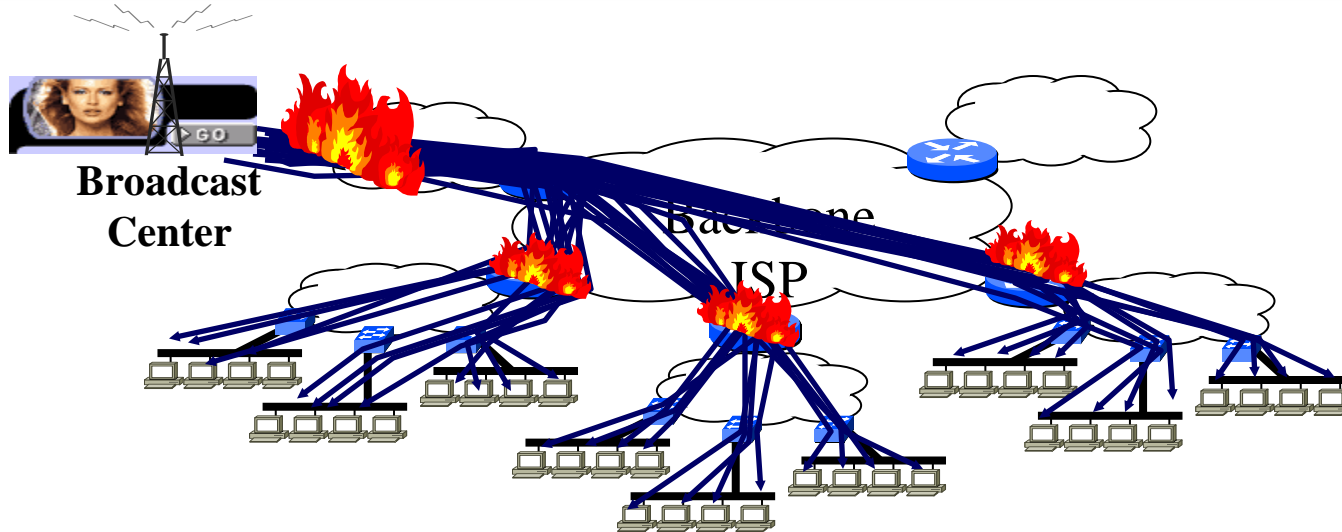
Ethernet Multicast

- Néhány Ethernet MAC cím multicast számára elkülönítve
- Ha egy G csoporthoz akarunk csatlakozni
 - A hálózati kártya (network interface card, NIC) elvileg csak a unicast és broadcast címre küldött csomagokat hallgatja
 - A csatlakozáshoz a G multicast címet is hallgatnia kell
 - Hardware megoldás, hatékony
- Csomagküldés a G csoportban
 - A csomag eláraszt minden LAN szegmenst
 - mint broadcast esetén
 - A kártyák melyek nem hallgatják a G multicast címet, eldobják a csomagot

Hálózati rétegű multicast

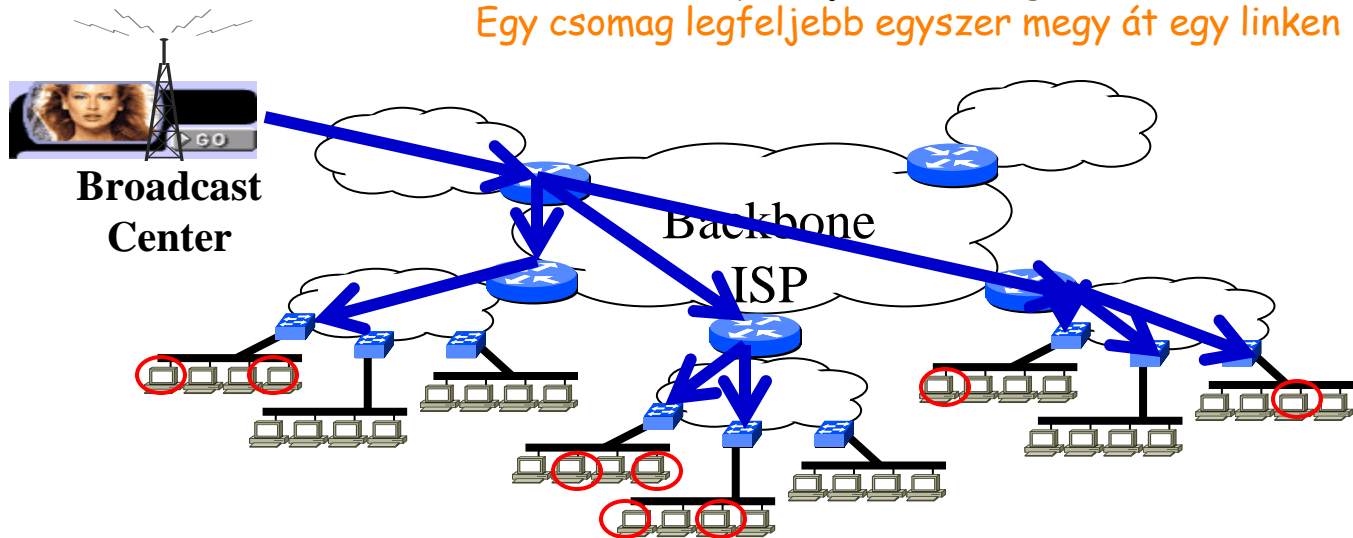
- A cél a hálózati erőforrások optimális kihasználása
 - **Egy csomag egy link-en csak egyszer megy át**
- A router-ek egy multicast fát tartanak fenn
 - A multicast fa mentén történik az adatátvitel
 - A router-ek duplikálják a csomagokat ha szükséges
 - Elágazási pontok a fán

A csoportos unicast nem skálázható

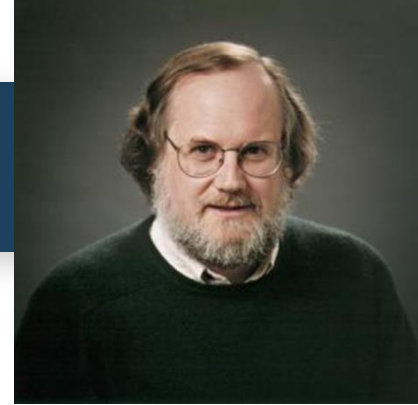


Helyette építsünk fákat

A router-ek duplikálják a csomagokat
Egy csomag legfeljebb egyszer megy át egy linken



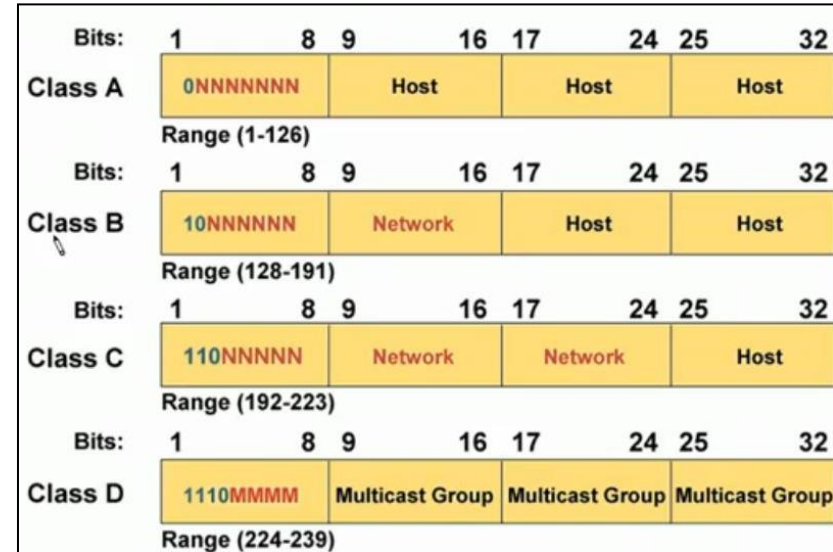
IP Multicast



- Steve Deering PhD disszertációja (1990)
 - **Any Source Multicast (ASM)**
- Nyitott csoportkommunikációs modell
 - Bárki csatlakozhat egy csoporthoz, bármilyen engedélyezés nélkül
 - Egy felhasználó több csoportnak is tagja lehet egyszerre
 - Bárki küldhet adatokat a csoportnak, ha nem is tagja annak
 - A csoport tagsága dinamikus
 - Senki nem ismeri a csoport méretét, vagy a tagok kilétét

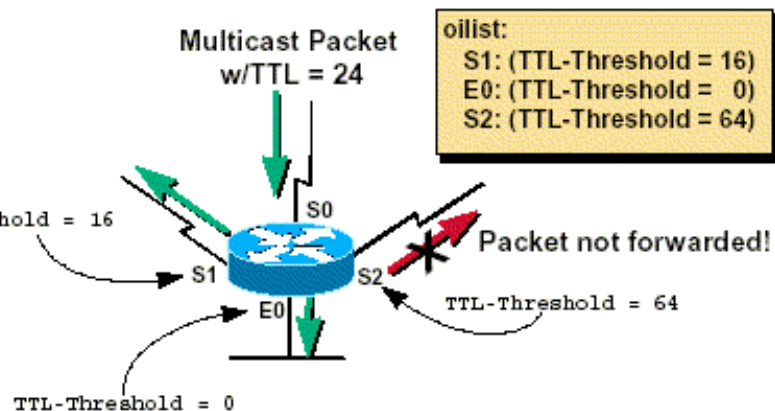
IP Multicast

- S. Deering, "Host Extensions for IP Multicasting", RFC 1112, 1989.
- A forrás csomagjait egy virtuális csoportcímre küldi
- Bárki aki csatlakozik a csoporthoz „elérhető” ezen a címen
 - Megkapja az ezen címre küldött csomagokat
- Egy multicast csoportot egy **class D** IP cím azonosít
 - **224.0.0.0 – 239.255.255.255**
 - 1110 + 28 bites csoport azonosító



Multicast Scoping

- Egy IP multicast csoport hatóköre szabályozva van:
 - TTL alapú szabályozás
 - Adminisztrációs szabályozás
- TTL alapú szabályozás
 - Node-local 0
 - Link-local 1
 - Site-local < 32
 - Region-local < 64
 - Continent-local < 128
 - Global Scope < 255



Multicast Scoping

- Adminisztrációs szabályozás
 - link-local scope 224.0.0.0 - 224.0.0.255
 - Egy router sohasem küldi tovább
 - global scope 224.0.1.0 - 238.255.255.255
 - A teljes Interneten érvényes
 - administrative scope 239.0.0.0 - 239.255.255.255
 - Nem küldik egy szervezet Intranet-jén kívülre

IP Multicast adatátvitel

- A csatlakozás egy multicast csoporthoz két lépésben történik
 - **A helyi hálózaton (LAN)**
 - Egy felhasználó értesíti a helyi multicast router-ét hogy szeretne csatlakozni egy csoporthoz
 - IGMP (IPv4), MLD (IPv6)
 - **A nagy kiterjedésű hálózaton (WAN)**
 - A helyi router közreműködik a hálózat többi multicast router-ével a multicast fa kiépítésében és a csomagok továbbításában
 - DVMRP, MOSPF, CBT, PIM-DM, PIM-SM, PIM-SSM

IGMP

- **Internet Group Management Protocol**
- IPv4 protokoll, a végső felhasználók és a helyi multicast router-ek között a helyi hálózaton
 - A multicast csoportokban való tagságot kezeli
 - Aszimmetrikus protokoll
 - Felhasználói rész
 - Router rész
- A router megtanulja hogy milyen csoportokat hallgatnak a saját helyi hálózatán
 - Nem érdekli hányan hallgatják, a fontos hogy legyen legalább egy valaki
 - Nem érdekli ki hallgatja

IGMPv1

- S. Deering, "Host Extensions for IP Multicasting", RFC 1112, 1989.
- A multicast router rendszeres **Query** üzeneteket küld az összes felhasználó közös multicast címére (224.0.0.1)
- A felhasználók **Report** üzenettel válaszolnak, melyben beszámolnak az általuk hallgatott csoportokról
 - A Report-ot a hallgatott csoportok multicast címére küldik
- A Report csomagok számának csökkentése érdekében:
 - Időzítők (timer) használata
 - Egy felhasználó nem válaszol azonnal
 - Host Suppression
 - Ha valaki más már válaszolt, törli a saját Report üzenetét
- **Unsolicited Report**
 - Ha egy felhasználó egy új csoportot akar hallgatni

IGMPv1 Router

- Egy IGMPv1 router fenntart egy multicast tagsági táblát
 - Milyen multicast csoportokat hallgatnak a hálózaton
 - Mikor volt az utolsó Report egy csoporttal kapcsolatban
- **Soft-state protokoll**
 - Ha egy adott időn belül nem erősíti meg senki egy csoport iránti érdeklődését, a csoportot törli a táblájából
- Csak azokat a multicast csomagokat küldi tovább a helyi hálózatra, melyeket egy a táblájában szereplő multicast címre küldtek

IGMPv2

- W. Fenner, "Internet Group Management Protocol, Version 2", RFC 2236, November 1997.
<http://www.ietf.org/rfc/rfc2236.txt>
- IPv6-os változata: MLD (Multicast Listener Discovery)
 - S. Deering, W. Fenner, B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, November 1999.
<http://www.ietf.org/rfc/rfc2710.txt>
- Bevezet egy gyors kilépési mechanizmust (**Fast Leave**)
 - Nem kell várni az időzítők lejáráásáig ahhoz, hogy a router „levágjon” egy csoportot

IGMPv2 üzenetek

- Membership Query
 - General Query vagy Group Specific Query
- Membership Report
- **Leave Group Message**
- Ha egy felhasználó ki akar lépni egy csoportból, küld egy Leave üzenetet az összes multicast router közös multicast címére (224.0.0.2)
- Mielőtt a router levágja a csoportot, megkérdezi, van-e valaki más aki hallgatja a csoportot
 - Group Specific Query
 - Ha egy adott időn belül nem érkezik válasz, a router törli a csoportot
- **IGMPv3** – később...

Multicast Routing

- A forrás egy multicast csoportcímre küldi csomagjait
- A hálózat multicast routerei kialakítanak és karbantartanak egy multicast fát
 - Az adatátvitel ezen fa mentén történik
- A helyi multicast router, az IGMP tagsági táblája alapján csatlakozik a fához, vagy elhagyja azt
- A hálózat multicast router-ei között egy útválasztó protokoll működik
 - MOSPF, DVMRP, CBT, PIM

MOSPF

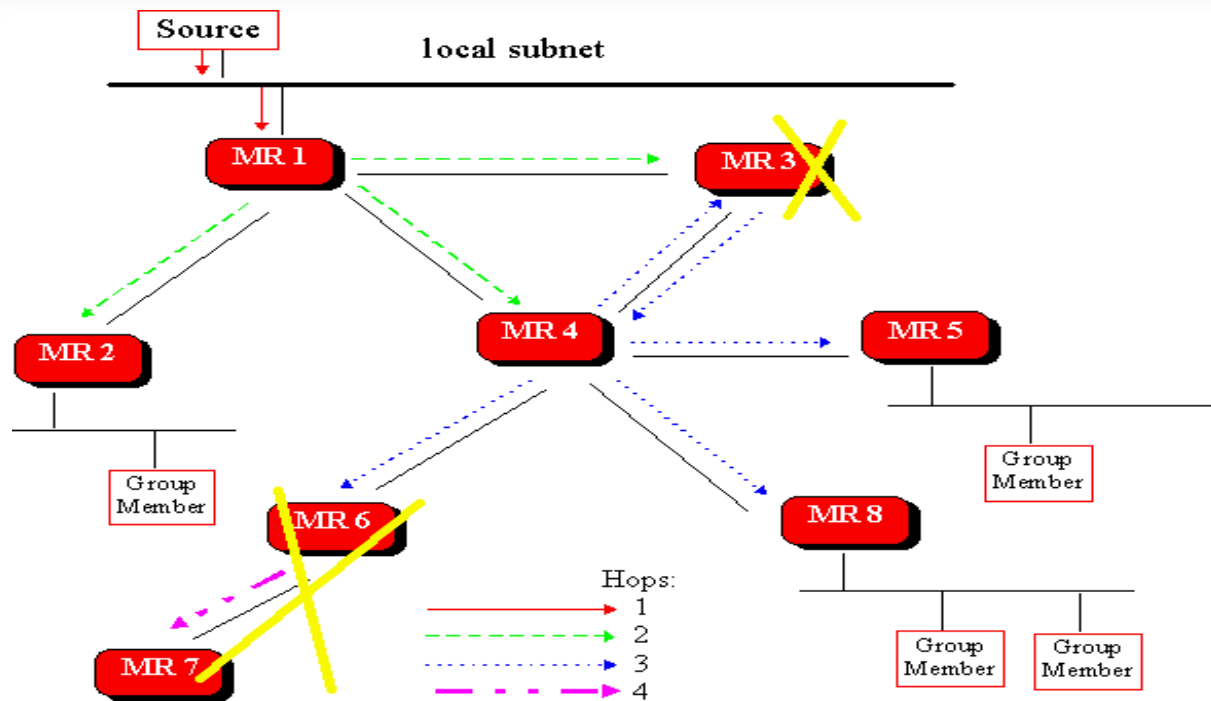
- **Multicast Open Shortest Path First**
 - J. Moy, „Multicast Extensions to OSPF”, RFC 1584, March 1994
<http://www.ietf.org/rfc/rfc1584.txt>
- Kapcsolatállapot (Link State) protokoll
- Az OSPF unicast útválasztó protokollt bővíti ki
 - Multicast csoportinformációt is küldenek a routerek egymásnak
 - Minden MOSPF router megtudja hogy melyik helyi hálózaton melyik csoportot hallgatják
 - Az információ alapján forrásonként és csoportonként egy legrövidebb útvonalú fát (shortest path tree) építenek fel
- Nagy a jelzés többletterhelés
- Nehezen alkalmazkodik a topológia változásokhoz
 - Újra kell számolni a fákat

DVMRP

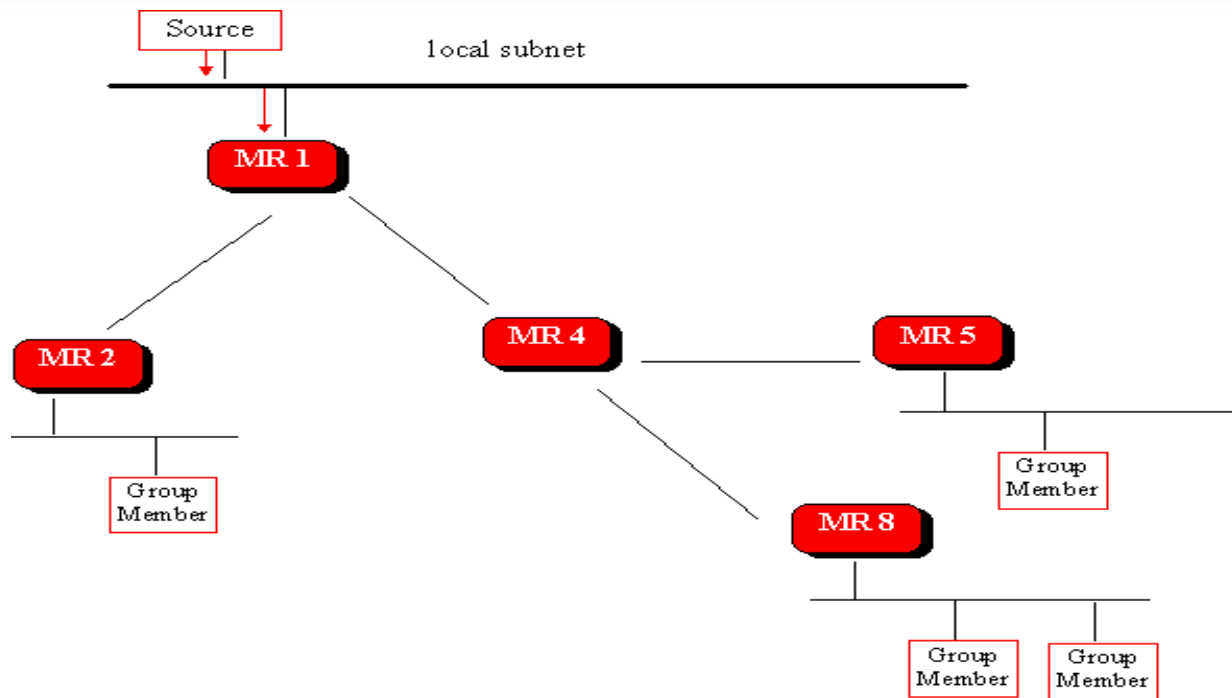
- **Distance Vector Multicast Routing Protocol**
 - D. Waitzman, C. Partridge, S. Deering, "Distance Vector Multicast Routing Protocol", RFC 1075, November 1988
<http://www.ietf.org/rfc/rfc1075.txt>
- Távolság-vektor (distance vector) alapú protokoll
 - A RIP unicast útválasztó protokollt használja

- **Elárasztás és metszés (flood and prune)**
 - **Elárasztás**
 - Ellenőrzi a csomag bejövő interfészét
 - Ha nem a legrövidebb út a forrás felé, eldobja a csomagot
 - [A RIP unicast routing tábla alapján](#)
 - Ha igen, továbbküldi a csomagot az összes többi interfészen
 - **Metszés**
 - Ha nincs érdekelt felhasználó egy helyi hálózaton
 - Ha nem a legrövidebb úton jött a csomag
 - **Egy közbeeső router megjegyzi azokat az interfészeit, ahol Prune érkezett**
 - Azokra az interfészekre nem küldi ki a további csomagokat
 - A Prune bejegyzések percenként elavulnak

DVMRP elárasztás



DVMRP metszés



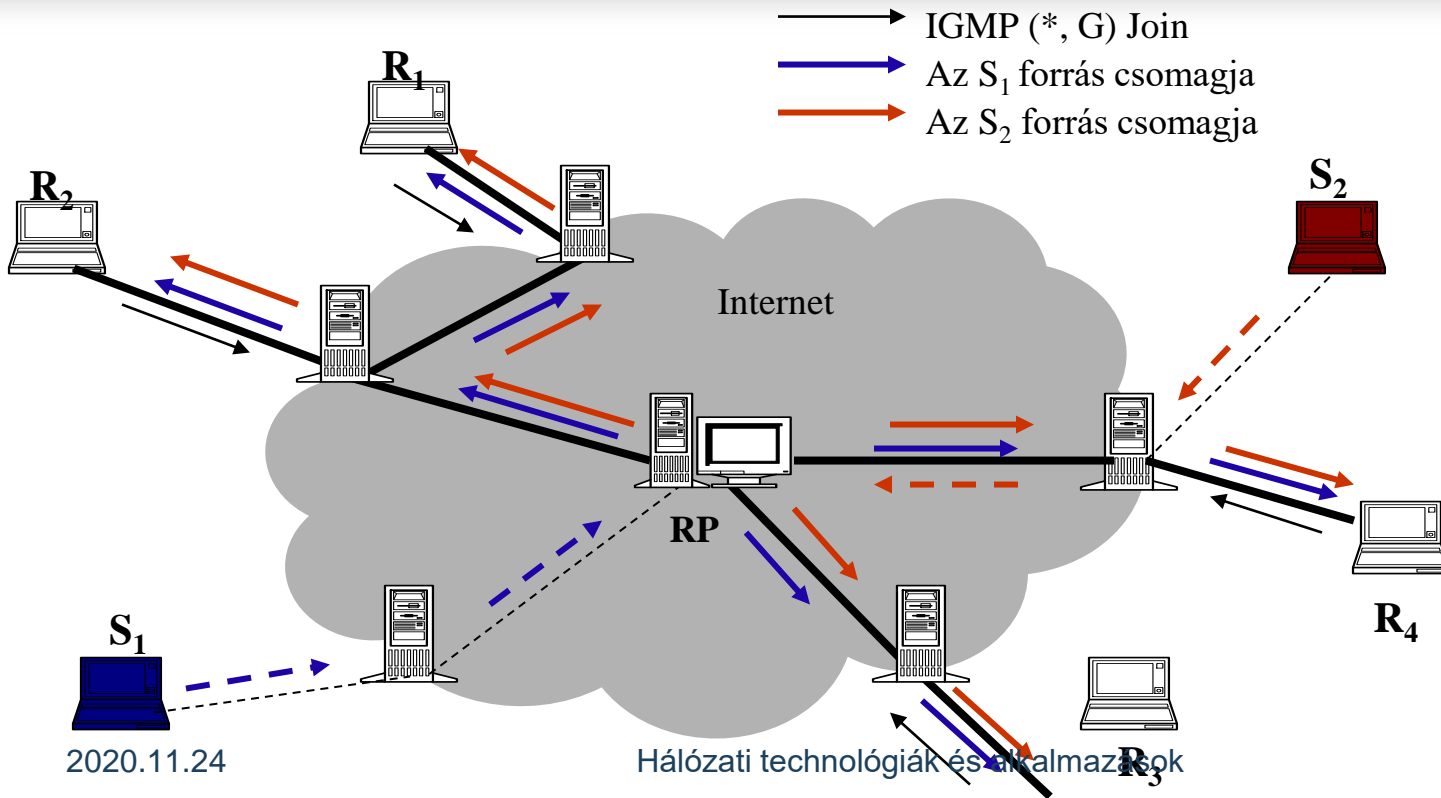
PIM

- **Protocol Independent Multicast**
 - PIM Dense Mode (PIM-DM)
 - PIM Sparse Mode (PIM-SM)
- PIM-SM
 - W. Fenner et al., „Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)”, RFC 4601, August 2006
 - Az egyik legelterjedtebb multicast útválasztó protokoll

PIM-SM

- Egy közös multicast fát használ (**shared tree**)
- Kiválaszt egy „randevú” pontot (RP)
 - Az RP a közös fa gyökere
 - Minden forrás az RP-hez küldi csomagjait
 - Az RP továbbküldi azokat a közös fán
 - Ha nagy forgalom, Source Register üzenettel bejelentkezik a forrás az RP-nél
 - az RP feliratkozik a forrás saját fájára
 - Bizonyos számú csomag után áttérés a közös fáról a forrás fájára
 - Nagy forgalmú forrásoknál az RP-nek csak közvetítő szerepe van, megismerteti a forrást a vevőkkel
 - Ha minden érdeklődő vevő már a forrás fáján van, az RP lecsatlakozik

PIM-SM csomagküldés



2020.11.24

Hálózati technológiák és alkalmazások

Az ASM modell hátrányai

- Az ASM modell elterjedését több gazdasági és technikai tényező gátolta
- **Bonyolult címkiosztás**
 - Dinamikus címválasztás a forrás által
 - Komplex címallokációs megoldások az ütközések elkerülésére
 - Multicast Address Allocation Architecture (RFC 2908)
 - MADCAP – Multicast Address Dynamic Client Allocation Protocol
 - AAP – Multicast Address Allocation Protocol
 - MASC – Multicast Address Set Claim

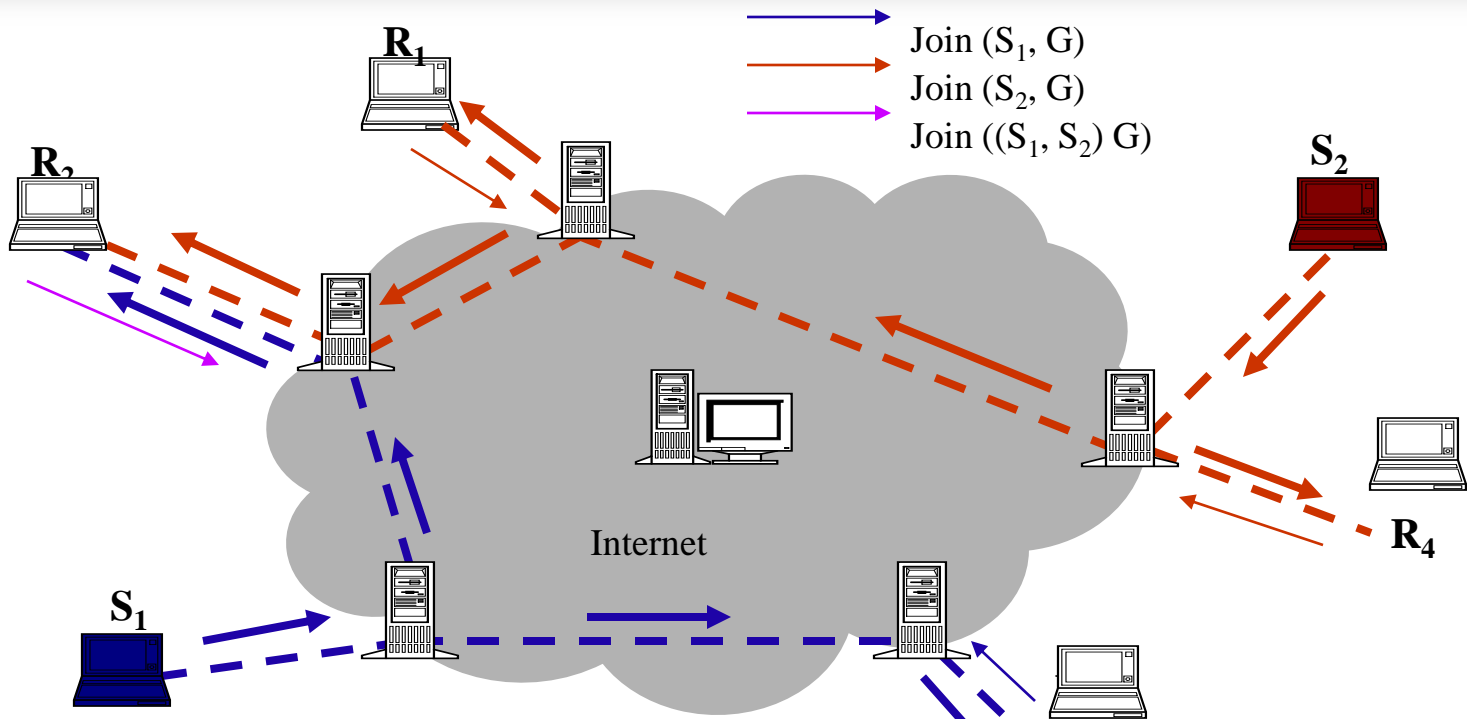
Az ASM modell hátrányai

- **Túl nyílt modell a szolgáltatók számára**
 - A források és vevők ellenőrizhetetlensége
 - Nehezen megoldható számlázás
- **Nem skálázható a tartományok közötti útválasztás**
 - PIM-SM csak egy tartományon belül
 - Egy ISP nem szereti ha forgalmát egy másik ISP-n belüli RP ellenőrzi
 - A tartományok között más protokollok
 - MSDP – Multicast Source Discovery Protocol
 - MBGP – Multicast Border Gateway Protocol

Az SSM modell

- Egy egyszerűbb modellre volt szükség
- **SSM - Source Specific Multicast**
 - Az Express modellre alapul
 - H. Holbrook, D. Cheriton, "IP Multicast Channels: Express Support for Large-Scale Single-Source Application", *ACM SIGCOMM'99*, Cambridge, MA, USA, Sept. 1999.
- A (*,G) multicast csoport helyett az (S,G) multicast csatornát használja
 - S a forrás unicast címe
 - G a csoport multicast címe
 - Csak az S forrás küldhet csomagokat az (S,G) csatorna vevőihöz
 - Az adatátvitel egy forrás-specifikus fa mentén történik

SSM csomagküldés



2020.11.24

Hálózati technológiák és alkalmazások

Forrás szűrés

- Az SSM-hez szükség van forrás szűrésre
 - A felhasználó nem csak azt mondja meg a helyi router-nek, hogy melyik csoportot hallgatja, hanem hogy azon belül melyik forrást is
- IPv4 – IGMPv3
 - B. Cain, et. Al, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.
<http://www.ietf.org/rfc/rfc3376.txt>
- IPv6 – MLDv2
 - R. Vida, L. Costa, „Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.
<http://www.ietf.org/rfc/rfc3810.txt>

Üzenet típusok

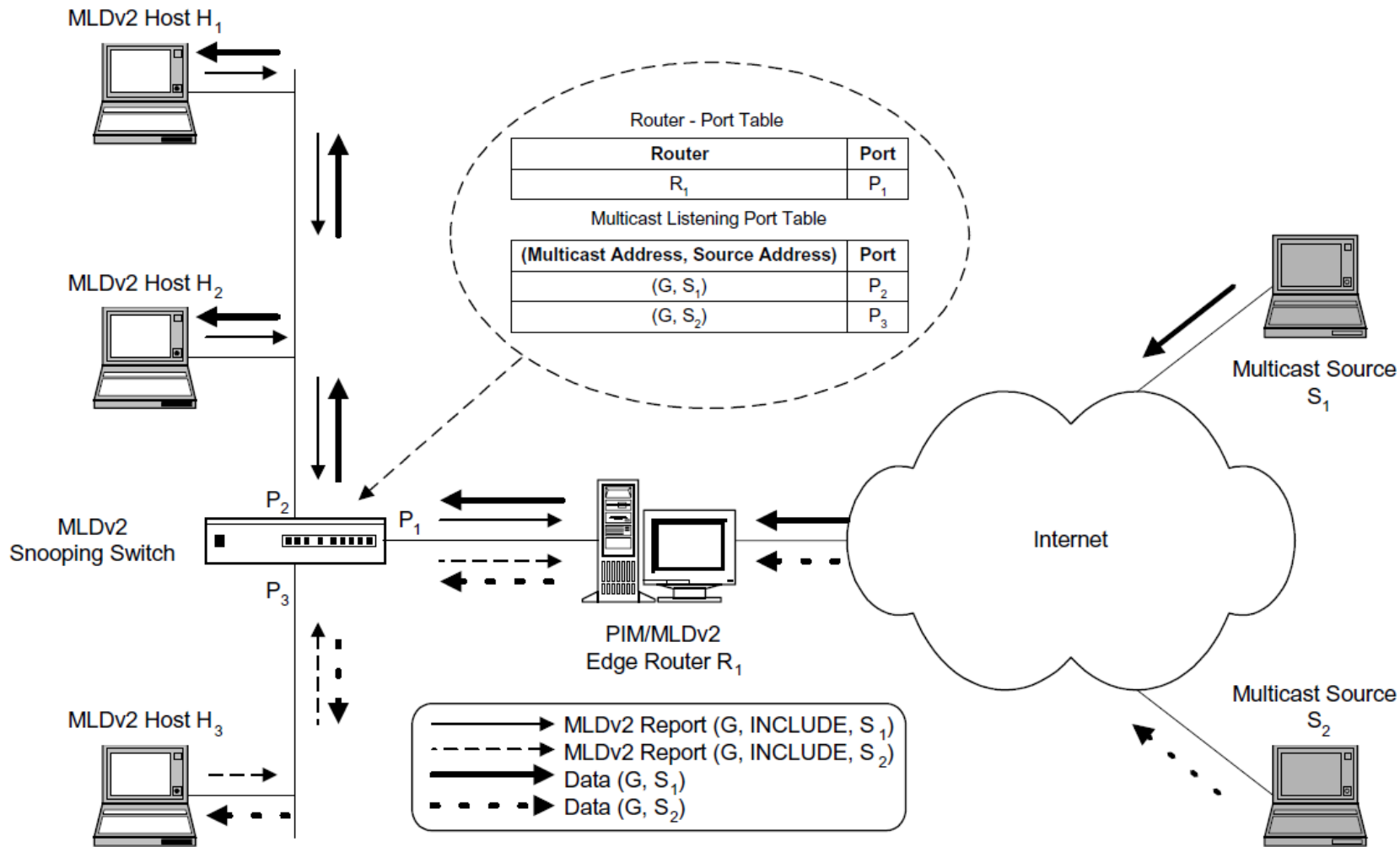
- IGMP/MLD **Query**
 - General Query
 - Ki mit hallgat?
 - Group Specific Query
 - Hallgatja-e valaki ezt a csoportot?
 - **Group and Source Specific Query**
 - Hallgatja-e valaki ezt a forrást ebben a csoportban?
- IGMP/MLD **Report**
 - Current State Record
 - Mit hallgatok – pl. Include (A) vagy Exclude (B)
 - A és B forráscím halmazok
 - Filter Mode Change Record
 - Szűrési mód váltása (Include vagy Exclude)
 - Source List Change Record
 - Allow (A) vagy Block (B)

Nincs host suppression!

- Túl komplex lenne

IGMP/MLD Snooping Switch

- A helyi hálózaton nem szeretnénk minden részt elárasztani a multicast forgalommal
 - Egy snooping switch kezeli ezt
- **Control message snooping**
 - Az IGMPv3 / MLDv2 Report üzeneteket csak a routerek felé küldjük
 - Mivel nincs host suppression, nem kell a többi vevő felé is küldeni
- **Multicast data snooping**
 - Csak oda küldjük az adatot, ahol vannak érdeklődők



IP Multicast

- Több éven át folyamatosan a jövő „forradalmi technológiájának” tartották
- **Előnyök**
 - Hatékony adatátvitel
 - A legrövidebb úton (DVMRP, MOSPF, PIM-SSM)
 - Figyelembe véve a fizikai topológiát
 - Hatékony erőforráskihasználás
 - Egy csomagot egy link-en csak egyszer küld át
 - Skálázható megoldás nagyméretű csoportok kommunikációjára
 - A csoportot egy virtuális cím azonosítja
 - Senki nem tartja számon a csoporttagok számát és kilétét

IP Multicast

- Mégsem terjedt el a várt mértékben
 - Technikai és gazdasági tényezők miatt
- **Technikai hátrányok**
 - Bonyolult címzés
 - Skálázható, tartományok közötti útválasztó megoldás hiánya
 - Rossz skálázhatóság a csoportok számát illetően
 - Egy router csoportonként egy bejegyzést tárol az útválasztó táblájában
 - A multicast címek nehezen aggregálhatók
 - Magasabb szintű szolgáltatások nehézkes támogatása
 - IP multicast egy *best-effort (több)pont-többpont* adatátviteli szolgáltatás
 - A végfelhasználók felelősek a felsőbb szintű szolgáltatások kezeléséért
 - Bonyolult torlódás vezérlés és megbízható adatátvitel

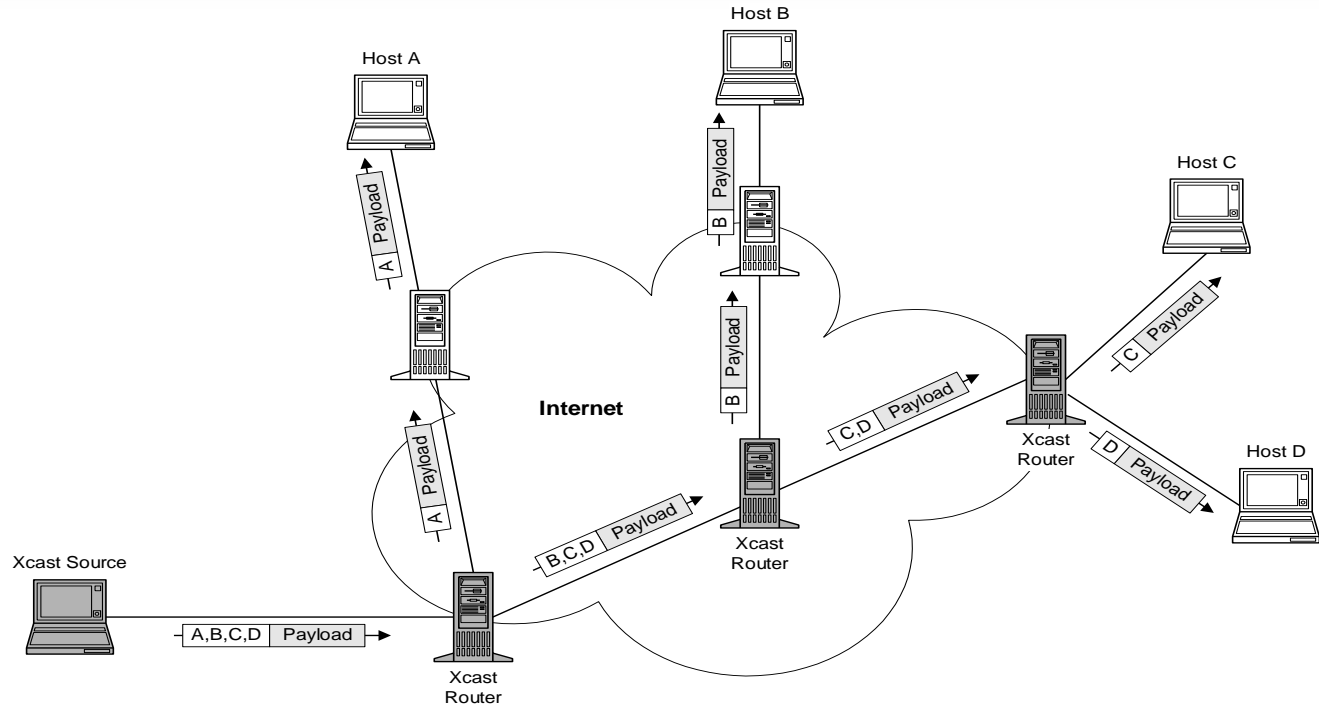
IP Multicast

- **Gazdasági tényezők**
 - Lassú és nehézkes telepítés a hálózatban
 - Noha a router-ek ma már képesek a multicast kezelésére, az ISP-k nem mindig aktiválják a hálózatukon
 - Csak akkor működik hatékonyan, ha minden router alkalmazza, különben alagutazni kell
 - „Tyúk-tojás” probléma
 - Az ISP-k nem támogatják, mert nincs elegendő multicast alkalmazás, nincs kellő kereslet
 - A szoftware cégek nem fejlesztenek multicast alkalmazásokat, mert nincs hálózati támogatás, nem lehet majd őket eladni
 - Nincs megfelelő gazdasági modell mögötte
 - Az ISP számára nehezen ellenőrizhető az erőforrásfelhasználás
 - A tartalom-szolgáltató számára nehezen ellenőrizhető ki használja a szolgáltatást
 - Nincs megfelelő számlázási megoldás

Explicit Multicast (Xcast)

- Hálózati rétegbeli multicast megoldás
- Nem használ multicast címezést
 - A forrás a csomag fejlécében tárolja az összes célállomás unicast IP címét
- A közbeeső Xcast router-ek duplikálással új csomagokat hoznak létre, a saját unicast útválasztó tábláik bejegyzései alapján
 - A router ellenőrzi hogy a kapott csomag fejlécében lévő célállomások felé melyik interfészen kell továbbítani a csomagot
 - Ennek megfelelően készíti el a duplikált csomagok fejléceit

Explicit Multicast (Xcast)



2020.11.24

Hálózati technológiák és alkalmazások

Explicit Multicast (Xcast)

- Nem skálázható megoldás nagy csoportokra
 - A fejlécbeli címek több helyet foglalhatnak el mint a tényleges adat
- Jól skálázható viszont nagyszámú kis csoport támogatására
 - A router-eknek nincs szükségük multicast útválasztó táblákra
- R. Boivie, N. Feldman, C. Metz, "Small Group Multicast: A New Solution for Multicasting on the Internet", *Internet Computing*, vol. 4, no. 3, May/June 2000, pp. 75-79.

Alternatív multicast megoldások

Van-e olyan csoportkommunikációs megoldás, mely ne szoruljon az ISP-k hálózati rétegbeli támogatására?

ALM – Application Layer Multicast

vagy...

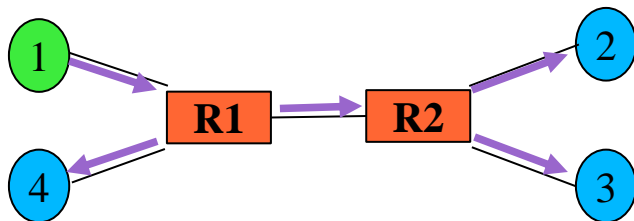
ESM – End System Multicast

vagy..

HBM – Host-based Multicast

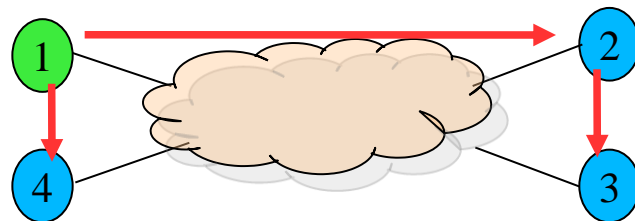
IP multicast - ALM

IP multicast



- Duplikálás a router-ekben
 - Hálózati támogatás
- A topológia függ...
 - az útválasztó tábláktól
 - a fizikai topológiától

ALM



- Duplikálás a végfelhasználóknál
 - Nem igényel hálózati támogatást
- Virtuális topológia
 - A fizikai topológia egy „fekete doboz”

ALM: motiváció

- Adatátvitel
 - Nem szükséges IP multicast támogatás
 - Kizárólag unicast kommunikációra épít
 - Kis csoportok
 - Az IP multicast nem mindig a legjobb megoldás
 - Az adatok aktív felhasználása
 - Az adatokat lehet módosítani/értékelni az átvitel folyamán
 - Az átviteli struktúra módosítható az adatok függvényében
- Kontroll
 - A kontroll adatok aggregálása (Megbízható multicast)

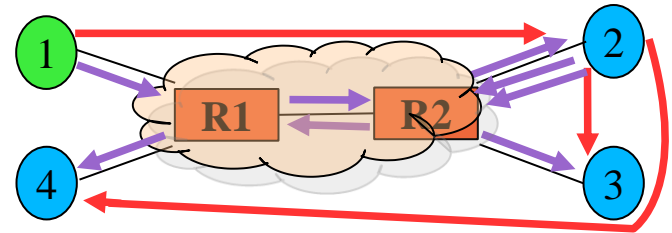
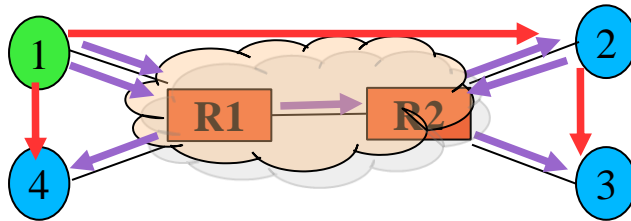
ALM: előnyök

- **Általános megoldás**
 - Nincs szükség hálózati támogatásra, bármilyen hálózat felett működhet
 - Felhasználhatóak a létező kommunikációs mechanizmusok
 - Pl. a TCP torlódás vezérlése
- **Skálázhatóság**
 - A router-ek nem tárolnak csoportonkénti bejegyzéseket
 - A peer-ek igen, de ők kevés csoportban vesznek részt
- **Egyszerűen telepíthető**
 - Nem szükséges a hálózat belső elemeit (router-ek) módosítani
 - Csak a végső felhasználónál kell telepíteni
- **Különböző metrikák a szomszédok kiválasztására**
 - Különböző topológiák
 - Nagyon precíz topológia kontroll lehetősége

ALM: hátrányok

Hatékonyság

- End-to-end “ágak”
 - A késleltetés nagyon nagy lehet
 - Nagy erőforrás (sávszélesség) pazarlás



Skálázhatóság

- Peer-ek közötti kapcsolatok folyamatos értékelése
 - Teljes gráf: $n \cdot (n-1)$ virtuális kapcsolat egy n tagú csoportban

ALM: hátrányok (2)

Stabilitás

- A résztvevők stabilitása
 - Az overlay hálózatban a résztvevők („router-ek”) a végfelhasználók
 - Kevésbé megbízhatóak mint egy valódi router
 - Jönnek-mennek a hálózatban
- A mérési adatok stabilitása
 - Az overlay hatékonysága a választott metrika stabilitásától is függ
 - RTT, sávszélesség, stb.
 - Mérlegelni kell a hatékony adatátvitelt a többletterhelés függvényében

Általános megjegyzések

- A legtöbb ALM megoldás megpróbálja kiküszöbölni az előbbi hátrányokat
 - Egy dolog biztos: **ALM soha nem lesz olyan hatékony mint az IP Multicast**
 - Különböző megoldások különböző kompromisszumokra alapulnak
- Mivel „multicast” kommunikációról van szó, a cél egy adatátviteli fa építése
 - A fa lehet szabályozott (maximalizálva a lehetséges gyerekek száma) vagy nem
 - A fa „minősége” attól függ, hogy mire akarjuk használni

ALM - Általános koncepció

- Az ALM megoldások két topológiába szervezik a résztvevőket
 - **Kontroll topológia („mesh” - szövevény)**
 - A kontroll topológia tagjai periódikus frissítő üzenetekkel ellenőrzik a „szomszédai” jelenlétét (hibák felderítése, kezelése)
 - **Adatátviteli topológia („tree” – fa)**
 - Az adatátviteli topológia a kontroll topológia egy része, melyet adatátvitelre használunk
- Ezen topológiák kiépítésének sorrendje alapján, az ALM megoldások lehetnek:
 - Mesh-first: Narada
 - Tree-first: Yoid, HMTP, TBCP, Overcast, ALMI
 - Implicit: CAN-Multicast, Scribe, Bayeux, NICE