

Számítógépes Hálózatok és Internet Eszközök

2011

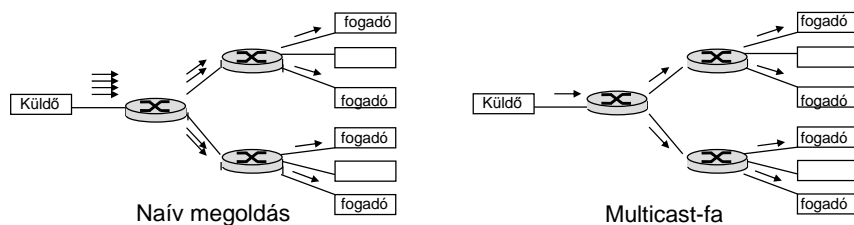
11. Hálózati réteg – Multicast

Broadcast és Multicast

- Broadcast routing
 - Egy csomagot (másolatot) minden más csomópontnak el kell küldeni
 - Megoldások:
 - A hálózat elárasztása (flooding)
 - Jobb: Konstruáljunk egy minimális feszítőfát
- Multicast routing
 - Az adatokat egy küldőtől egyidejűleg több fogadóhoz kell eljuttatni
 - Real time Streaming,
 - Web-cache update,
 - IPTV,
 - Multi-player játékok
 - Telefon-, Videokonferencia (all-to-all multicast),
 - ...

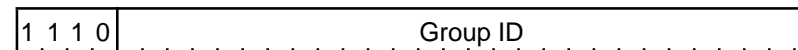
Multicasting

- Naív megoldás: Multicast-via-Unicast:
 - A küldő egy külön másolatot küld az adatokról minden fogadónak.
 - Nagyon inefficiens: A küldött csomagok száma sokkal nagyobb, mint ami szükséges lenne (különösen rossz all-to-all multicast esetén).
- Egy multicast-fa felepítése segítségével:
 - Minden linken csak egyszer továbbítódik egy csomag.
 - A routerek döntenek el, hogy egy csomagot több linken is továbbítanak-e.



IP Multicast Címek

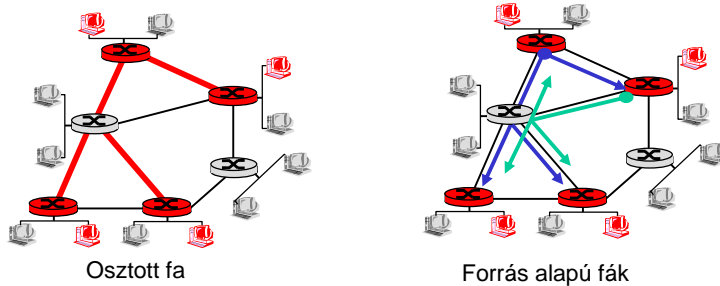
- Class D IP címek 224.0.0.0 – 239.255.255.255
 - Egy multicast-csoport (group) minden tagja ugyanazt a címet használja



- Cím allokálása
 - Jól ismert (well-known) multicast címeket az IANA (Internet Assigned Numbers Authority) adja ki
 - Átmeneti (Transient) multicast címek dinamikusan allokálhatók és visszadhatók
- Hatralévő rész: Jim Kurose (UMASS) foliái alapján

Multicast Routing: A megoldandó probléma

- **Cél:** találjunk egy fát (vagy fákat), amely összeköti azokat a routereket, melyek lokális hálózatában van olyan állomás, amely tagja a multicast csoportnak
 - **fa:** nem tartalmaz kört (nem minden utat használ a routerek között)
 - **forrás alapú (source-based):** minden forrástól külön fa a fogadókhoz
 - **osztott fa (shared tree):** ugyanazt a fát használja a csoport minden tagja



Módszerek multicast fák felépítésére

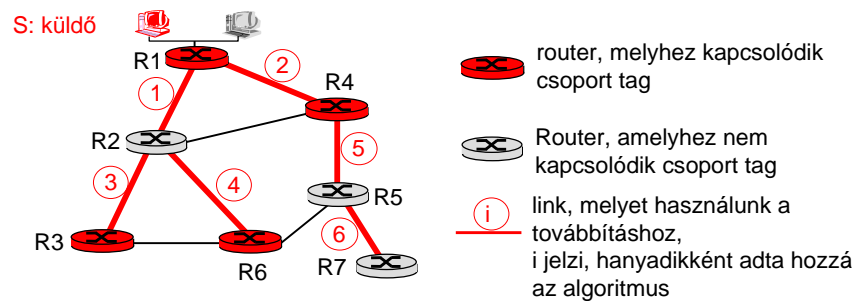
Módszerek:

- **source-based tree (forrás alapú fa):** küldőnként egy fa
 - legrövidebb utak fája
 - utak megfordítása (reverse path forwarding)
- **group-shared tree:** a csoport ugyanazt a fát használja
 - Steiner fa
 - center-based trees

...először áttekintjük az alap módszereket, majd a specifikus protollokat, hogy hogyan használják ezeket a módszereket

Legrövidebb utak fája (Shortest Path Tree)

- multicast forwarding tree: a legrövidebb utak a küldőtől minden fogadóhoz
 - Dijkstra algoritmus

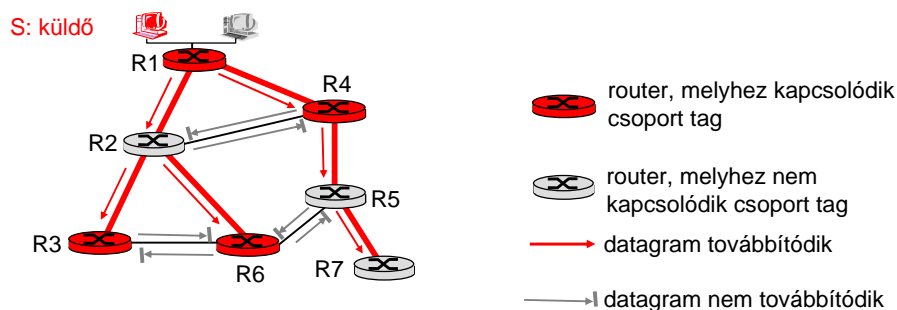


Utak megfordítása (Reverse Path Forwarding)

- Azon alapul, hogy a routerek tudják a legrövidebb utat a küldőhöz
- Minden router a következő szabály szerint továbbítja a csomagot:

```
if (multicast datagramot fogadunk a küldő felől a legrövidebb úton a bejövő linken)
then minden kimenő linken továbbítjuk a datagramot (elárasztás)
else ignoráljuk a datagramot
```

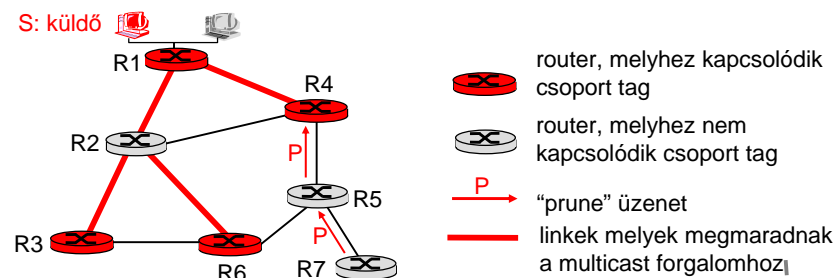
Reverse Path Forwarding: példa



- Az eredmény egy küldő-specifikus fordított legrövidebb utak fája
 - Aszimmetrikus linkek esetén rossz választás lehet

Reverse Path Forwarding: A fa “visszavágása” (pruning)

- A fa tartalmazhat részfákat, melyekben egyetlen egy multicast csoport tag sincs. Ilyen részfákba nem kell a datagramokat továbbítani
 - “prune” üzeneteket küldenek visszafelé azok a routerek, amelyekhez tartozó részfában nincs csoport tag



Shared-Tree: Steiner fa

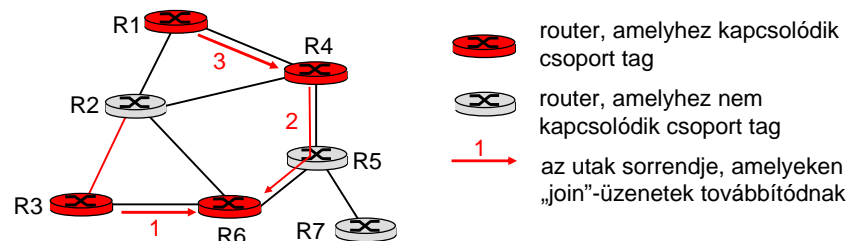
- Steiner fa:** egy fa, amely minden routert összeköt, amelyhez kapcsolódik multicast csoport tag
- Minimum Steiner fa:** minimális költségű fa, amely minden routert összeköt, amelyhez kapcsolódik multicast csoport tag
- A minimum Steiner fa kiszámítása NP-teljes probléma
- kiváló heurisztikák léteznek
- a gyakorlatban nem használják:
 - számítási bonyolultság
 - a teljes hálózatról szükséges információ
 - monolitikus: újraszámítás, ha egy új router csatlakozik a hálózathoz, vagy egy elhagyja a hálózatot

Centrum alapú fák (Center-based trees)

- Egy közös fát használ mindenki
- Egy routert azonosítunk a fa **centrum**aként
- Kapcsolódás a fához:
 - A csatlakozni kívánó router küld egy unicast *join*-üzenetet a centrum-routernek
 - A „join”-üzenetet közbenső routerek továbbítják a centrum-routernek
 - A „join”-üzenet elér egy olyan routert, amely már benne van a fában (legrosszabb esetben a centrum-router)
 - Az út, amin a „join”-üzenet továbbítódott, egy új ága lesz a fának, amely az elért routerből ágazik le

Centrum alapú fák: egy példa

Tegyük fel, hogy R6 lett centrumnak választva:



Internet Multicasting Routing: DVMRP

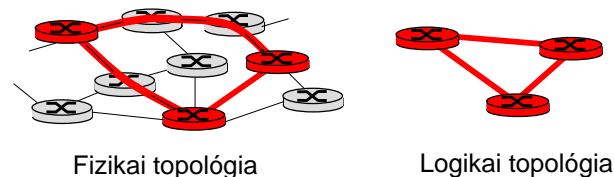
- **DVMRP**: distance vector multicast routing protocol, RFC1075
- **flood-and-prune (elárasztás és visszavágás)**: reverse path forwarding (RPF), source-based tree
 - Egymással kommunikáló DVMRP routerek felépítenek egy RPF fát, amelyet a DVMRP saját routing táblájában tárol
 - Semmit sem teszünk fel a DVMRP routerek között felhasznált unicastról
 - A multicast csoportot inicializáló datagram mindenfelé továbbítódik RPF-elárasztást használva
 - Azok a routerek, melyek nem kívánnak csatlakozni a fához, „prune”-üzenetet küldenek vissza

DVMRP: folytatás...

- **soft state**: DVMRP router periodikusan (1 percenként) “elfelejti” a visszavágott (pruned) ágakat:
 - A multicast adatok továbbfolyznak a nemvisszavágott ágba
 - A router egy ilyen ágban vagy megint visszaküld egy „prune”-üzenetet vagy tovább fogadja az adatokat
- Routerek gyorsan tudnak csatlakozni a fához
 - Miután egy szomszédos állomástól IGMP „join”-üzenetet kaptak
- Általánosan implementált kommerciális routerekben
- Mbone routing DVMRP-t használ

Alagút (Tunneling)

Hogy lehet összekötni a multicast routereket az unicast routerek „tengerében”?



- A multicast datagramot beágyazzuk egy „normális” (nem-multicast-címzésű) datagram belsejébe
- A „normális” IP datagram-ot rendes IP unicast-ot használva „egy alagúton” küldjük a fogadó multicast routerhez
- A fogadó multicast router „kiágyazza” multicast datagramot az unicast datagram belsejéből

PIM: Protocol Independent Multicast

- Nem függ semmilyen specifikus unicast routing algoritmustól (mindegyikkel működik)
- Két különböző multicast scénárió :

Sűrű:

- A csoport tagjai sűrűn helyezkednek el egymáshoz közel
- Sávzélesség bőven rendelkezésre áll

Ritka:

- A hálózatok száma, melyek tartalmaznak csoport tagot, összehasonlítva a kapcsolódó hálózatok számával, alacsony
- A csoport tagjai nagy távolságra szétszórtak
- Sávzélesség nem áll olyan bőségesen rendelkezésre

A sűrűség / ritkaság következményei:

Sűrű:

- A routereknél feltételezzük a csoport tagságot, addig amíg a router explicit „prune” üzenetet küld
- A multicast fa konstrukciója *adatok által irányított* (pl. RPF)
- Sávzélességet és a nem-csoport-routerek erőforrását pazarolja

Ritka:

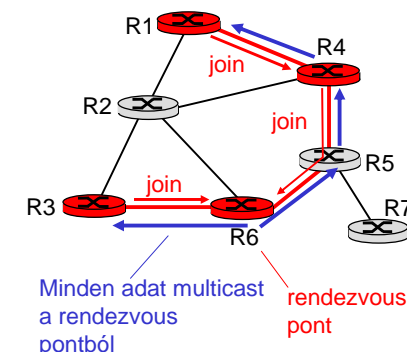
- Amíg a router nem kapcsolódik explicit, addig nincs csoport tagság
- A multicast fa konstrukciója *fogadó által irányított* (pl., centrum alapú fa)
- Sávzélességet és a nem-csoport-routerek erőforrását konzervatíván használja

PIM Sűrű Mód (PIM Dense Mode)

- **flood-and-prune RPF**, hasonló a DVMRP-hez, de
- A felhasznált unicast protokoll routing tábláját használja az RPF-hez
- Kevésbé bonyolult (kevésbé hatékony) elárasztás lefelé, mint amit DVMRP használ, csökkenti a függést a felhasznált routing algoritmustól
- tartalmaz olyan protokoll mechanizmusokat, amellyel egy router kideríti, hogy levél csomópont-e a fában

PIM – Ritka Mód (PIM Sparse Mode)

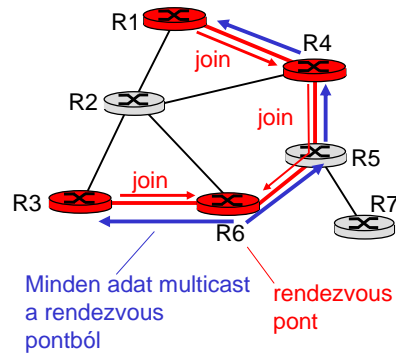
- Centrum alapú fa felépítése
- A kapcsolódó router „join” üzenetet küld a rendezvous ponthoz (RP)
 - Közbenső routerek aktualizálják az állapotukat és továbbítják a „join” üzenetet
- Miután az RP által kapcsolódott a router, a router átválthat forrás alapú fára
 - Növekszik az átvitel: kevésbé centralizált, rövidebb utak



PIM – Ritka Mód

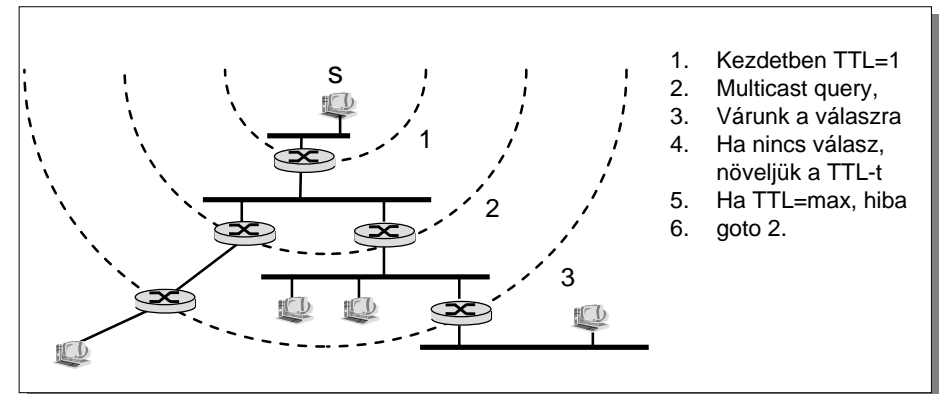
Küldő(s):

- Az adatot unicasttal küldi a RP-nak, ami lefelé elosztja a fában, amelynek az RP a centruma
- A RP kibővítheti a multicast fát a forráshoz
- A RP küldhet „stop” üzenetet, ha nincs fogadó
 - “senki se hallgat!”



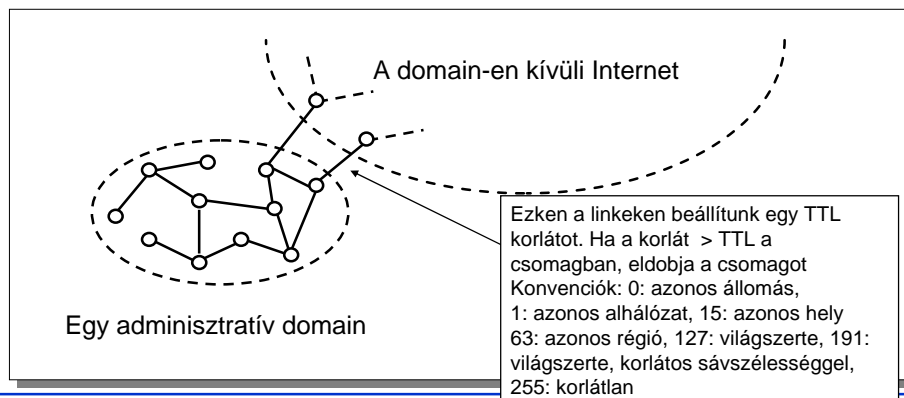
Multicast Scope Control – kicsi TTL esetén

- Táguló gyűrű keresés (expanding ring search) a csoport egy közeli részalmazának megtalálásához



Multicast Scope Control – Nagy TTL esetén

- Adminisztratív TTL korlát a multicast forgalmat az adminisztratív domain-en belül tartja, pl. biztonsági vagy erőforrás okokból



Multicast Scope Control

- Administratively-Scoped Addresses (RFC 1112)
 - 239.0.0.0 — 239.255.255.255 közti címekeket használ
 - Támogat átfedő (nem csak egymásba ágyazott) domain-eket is

