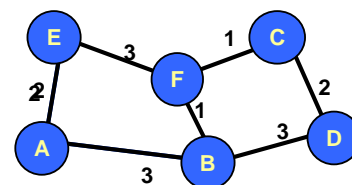


# Számítógépes Hálózatok 2011

## 9. Hálózati réteg – Distance Vector Routing, Link State Routing, RIP, IGRP, OSPF, Inter-AS Routing, BGP

### Distance Vector Routing Protokoll

- A Bellman-Ford algoritmusnak az **elosztott** változatát használja, azaz minden csomópont csak a direkt szomszédjaival kommunikál
- **Asszinkron** működés
  - A csomópontoknak nem ugyanabban a „körben” kell információkat cserélniük
- Minden router nyilvántart egy táblát minden lehetséges célhoz egy bejegyzéssel (**distance vector**)
  - egy bejegyzés tartalmazza
    - a legrövidebb út (becsült) költségét (delay, vagy #hops)
    - a következő csomópont címét ezen az úton (next hop)
- minden router ismeri a költséget a direkt szomszédjaihoz
- Periodikusan elküldi a tábláját minden szomszédjának
- Amikor egy router megkapja a szomszéd tábláját aktualizálja a saját tábláját



Initial distance vector of A

A	cost	next hop
B	3	B
C	$\infty$	-
D	$\infty$	-
E	2	E
F	$\infty$	-

Initial distance vector of B

B	cost	next hop
A	3	A
C	$\infty$	-
D	3	D
E	$\infty$	-
F	1	F

A's vector after A received B's vector

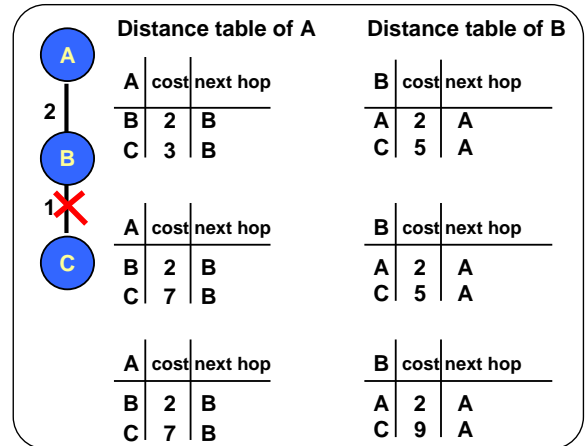
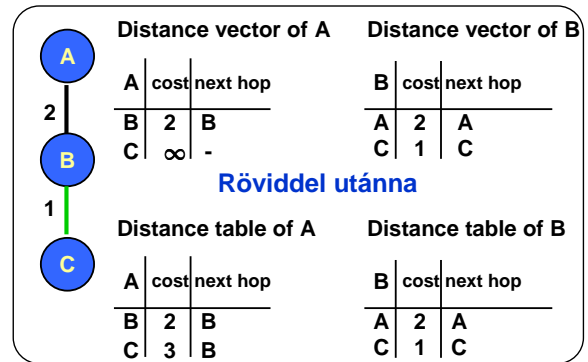
A	cost	next hop
B	3	B
C	$\infty$	-
D	6	B
E	2	E
F	4	B

A's final distance vector

A	cost	next hop
B	3	B
C	5	B
D	6	B
E	2	E
F	4	B

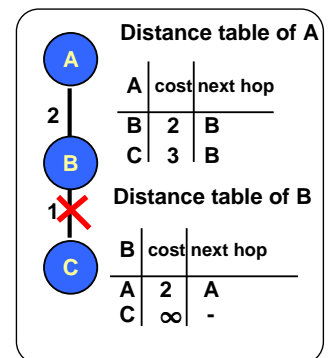
## “Count to Infinity” Probléma

- „Jó hír” gyorsan terjed
  - Új kapcsolat létrejöttkor gyorsan aktualizálódnak a táblák
- „Rossz hír” lassan terjed
  - Kapcsolat kiesik
  - A szomszédok felváltva növelik a távolságokat
  - “Count to Infinity” Probléma
    - A és B nem tudja, hogy C nem elérhető (amíg a távolság el nem ér egy limitet, amit  $\infty$ -nek tekintenek)
  - Ciklusok keletkezhetnek



## “Count to Infinity” Probléma

- Módosítások a Distance-Vector routing protollokban
  - a ping-pong-ciklusokat (count to infinity) megakadályozásához
  - **split horizon**: olyan utakat nem küld vissza a csomópont annak a szomszédjának, amit tőle „tanult”
    - a példában A nem küldi a (C,3,B) sornak megfelelő utat vissza B-nek, mert azt B-től kellett „tanulnia”
  - **split horizon with poison reverse**: negatív információt küld vissza
    - A pl. (C, $\infty$ ) utat küldi vissza B-nek
- Mindkét módszer csak két csomópontból álló ciklust kerül el



## Link State Protokoll

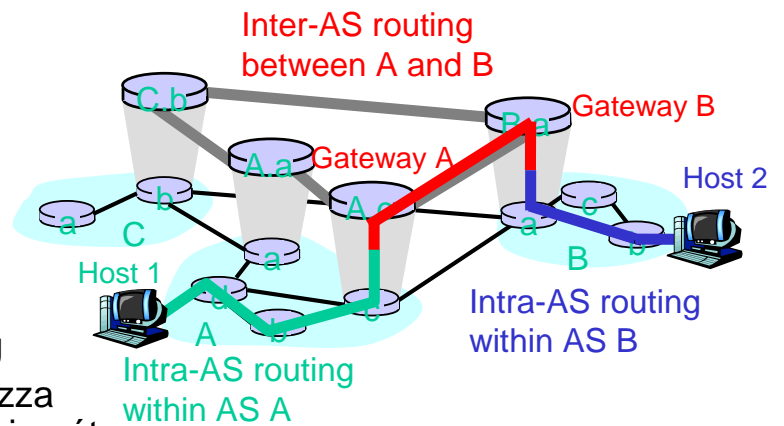
- Minden Link State router
  - tárolja a hálózat topológiáját
  - egy nem-elosztott legrövidebb utak algoritmust használ
- A routerek **Link State Packets (LSP)** által cserélnék ki információkat
- LSP tartalmazza
  - az LSP-t létrehozó r router IP címét
  - a költségét r minden direkt szomszédjához
  - sorozatszámot (SEQNO)
  - TTL (time to live) mezőt
- Megbízható elárasztás (Reliable Flooding)
  - minden csomópont aktuális LSP-jét tároljuk
  - továbbítjuk az LSP-eket minden szomszédos csomóponthoz
    - azon csomópont kivételével, amely az LSP-t felénk továbbította
  - A továbbításnál csökkentjük a TTL értékét
  - periodikusan létrehozunk egy új saját LSP-t
    - növekvő SEQNO-val

## A „lapos” routing korlátai

- Link State Routing
  - $O(D \cdot n)$  bejegyzésre van szükség, ahol  $n$  a routerek száma,  $D$  a maximális fok
  - Minden csomópont minden más csomópontnak el kell hogy küldje az információit
- Distance Vector
  - $O(n)$  bejegyzés routerenként
  - Ciklusokat okozhat
  - Konvergencia ideje a hálózat méretével nő
- Az Internet több mint  $10^6$  routert tartalmaz
  - ezek a u.n. „lapos” routing módszerek nem használhatók az egész Internetre
- Megoldás:
  - Hierarchikus routing

## Autonomous Systems (AS), Intra-AS és Inter-AS routing

- Autonomous Systems (AS)
  - Egy két szintű modellt ad a routinghoz az Interneten
  - Példa AS-re: elte.hu
- Intra-AS-routing
  - routing az AS-en belül
  - pl. RIP, OSPF, IGRP, ...
- Inter-AS-routing
  - Kapcsolódási pont: **átjáró (gateway)**
  - teljesen decentralis routing
  - Mindeki saját maga határozza meg az optimalizálási kritériumát
  - pl. BGP, EGP (korábban)



## Intra-AS routing: RIP Routing Information Protocol (RFC 1058)

- Distance Vector algoritmus
  - távolság metrika = hop szám (linkek száma)
- A távolság vektorokat (distance vector) minden router minden 30s Response-üzenettel (advertisement) adja át a szomszédjának
- A szomszédok szintén egy új advertisement-et küldenek ha a táblájuk ezáltal megváltozott
- Minden Advertisement-ben
  - célhálózathoz hirdetik meg az utakat UDP-vel (UDP port 520)
- Ha 180s-ig nem kap a router advertisement-et egy szomszédjától
  - az utakat a szomszédon keresztül érvénytelennek deklarálja
  - új Advertisement-eket küld a szomszédainak
- Hogy elkerülje a ping-pong-ciklusokat (count to infinity), „split horizon with poison reverse” módszert használ
- Végtelen távolság = 16 Hop (limitet szab a hálózat átmérőjére)

## Intra-AS routing: OSPF routing (Open Shortest Path First)

- “open” = nyilvánosan rendelkezésre álló
- Link-State algoritmus
  - LS csomagok terjesztése
  - a topológiát minden csomópontban tárolja
  - az útvonalakat Dijkstra algoritmusával számítja ki
- OSPF-advertisement
  - TCP-vel, növeli a biztonságot (security)
  - az egész AS-be elárasztja (broadcast)
  - több egyenlő költségű útvonal lehetséges

## Intra-AS routing -- Hierarchikus OSPF

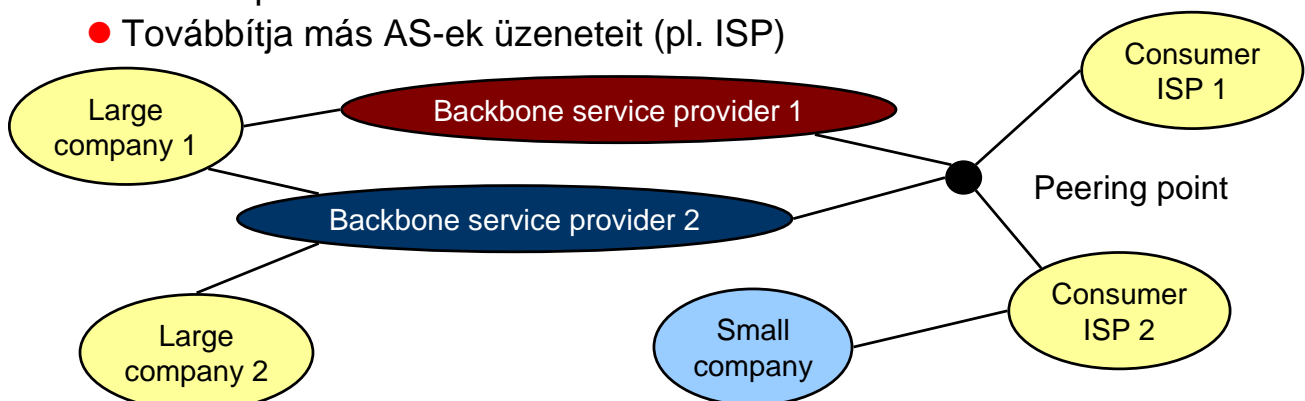
- Nagy hálózatokhoz két hierarchia szint:
  - Lokális terület és gerinchálózat (backbone)
    - Lokális: Link-state advertisement
    - Minden csomópont csak az irányt számítja ki más lokális területek hálózataihoz
- Local Area Border Router:
  - A saját lokális területeik távolságait foglalják össze
  - Ezeket más Lokal Area Border Router-eknek meghirdetik (advertisement)
- Backbone Routers
  - OSPF protokollt használnak a gerinchálózatra korlátozva
- Boundary Routers:
  - Más AS-ekkel kapcsolnak össze

## Intra-AS routing: IGRP (Interior Gateway Routing Protocol)

- CISCO-Protokoll (1980-as évek közepe), a RIP utódja
- Distance-Vector-Protokoll, mint a RIP
  - Holddown time
  - Split horizon
  - Poison reverse
- Különböző költség metrikákat támogat
  - Delay, Bandwidth, Reliability, Load, stb...
- TCP-t használ a routing információk kicseréléséhez

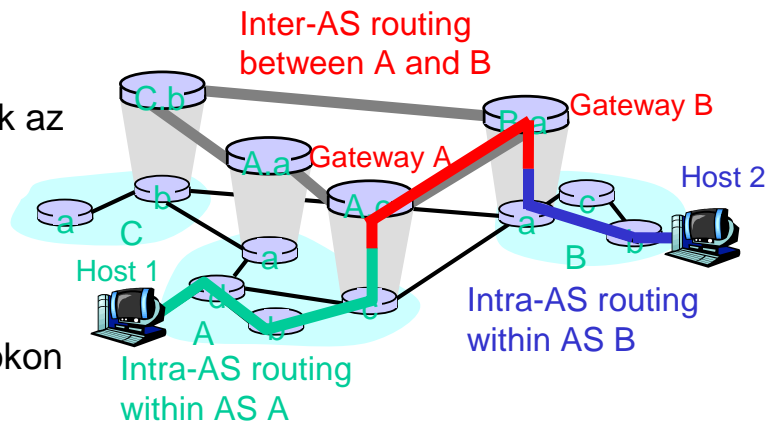
## Autonóm rendszerek (AS) típusai

- Stub-AS
  - Csak egy más AS-hez kapcsolódik
- Multihomed AS
  - Több AS-hez kapcsolódik
  - Nem továbbítja más AS-ek forgalmát
- Transit AS
  - Több kapcsolat
  - Továbbítja más AS-ek üzeneteit (pl. ISP)



## Inter-AS-Routing

- Inter-AS-Routing nehéz...
  - Szervezetek megtagadhatják az üzenetek továbbítását (pl. csak fizető ügyfelek csomagjait továbbítja)
  - Politikai követelmények
    - Továbbítás más országokon keresztül?
  - Különböző AS-ek routing-metrikái sokszor nem összehasonlíthatók
    - Útvonal optimalizálás lehetetlen!
    - Inter-AS-Routing megpróbálja legalább a csomópontok elérhetőségét lehetővé tenni
  - Méret: inter-domain routereknek ma kb. 140.000 hálózatról kell tudni



## Inter-AS routing: BGP (Border Gateway Protocol)

- Az inter-AS routing standard BGPv4
- Path Vector protokoll
  - Hasonló a Distance Vector protokollhoz
  - Minden Border Gateway meghirdeti minden szomszédjának (peers) az egész utat (AS-ek sorozata) a célig (advertisement)
  - TCP-t használ
- Amikor Gateway X az utat Z-hez Peer-Gateway W-nek küldi
  - akkor W választhatja ezt az utat, vagy éppen nem
  - Optimalizálási kritériumok:
    - költségek, politika, etc...
  - Ha W az X által meghirdetett utat választja, akkor meghirdeti
    - $\text{Path}(W,Z) = (W, \text{Path}(X,Z))$
- Megjegyzés
  - X tudja szabályozni a hozzá érkező forgalmat a meghirdetések által.
  - Komplikált protokoll

## 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, Video-On-Demand,
    - Telefon-, Videokonferencia (all-to-all multicast),...
  - IP D címosztály:
    - Egy multicast-csoport (group) minden tagja ugyanazt a címet használja
  - Megoldások:
    - Optimális: Minimális Steiner Fa Probléma
      - NP-teljes (2-approximáció  $O(n \log n)$  idő alatt kiszámítható!)
    - Más (nem-optimális) fát konstruálni

## Multicasting

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

