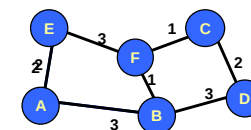


Számítógépes Hálózatok 2010

10. Hálózati réteg – RIP, IGRP, OSPF, Inter-AS Routing, BGP, IP címzés, ARP, DNS, Circuit Switching, Packet Switching

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 szomszédaihoz
- Periodikusan elküldi a tábláját minden szomszédjának
- Amikor egy router megkapja a szomszéd tábláját aktualizálja a saját tábláját



Initial distance vector of A			Initial distance vector of B		
A	cost	next hop	B	cost	next hop
B	3	B	A	3	A
C	∞	-	C	∞	-
D	∞	-	D	3	D
E	2	E	E	∞	-
F	∞	-	F	1	F

A's vector after A received B's vector			A's final distance vector		
A	cost	next hop	A	cost	next hop
B	3	B	B	3	B
C	∞	-	C	5	B
D	6	B	D	6	B
E	2	E	E	2	E
F	4	B	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 ∞ -nek tekintenek)
 - Ciklusok keletkezhetnek

Distance vector of A			Distance vector of B		
A	cost	next hop	B	cost	next hop
B	2	B	A	2	A
C	∞	-	C	1	C

Röviddel utána

Distance table of A			Distance table of B		
A	cost	next hop	B	cost	next hop
B	2	B	A	2	A
C	3	B	C	1	C

Distance table of A			Distance table of B		
A	cost	next hop	B	cost	next hop
B	2	B	A	2	A
C	3	B	C	5	A

1X

Distance table of A			Distance table of B		
A	cost	next hop	B	cost	next hop
B	2	B	A	2	A
C	7	B	C	5	A

A

Distance table of A			Distance table of B		
A	cost	next hop	B	cost	next hop
B	2	B	A	2	A
C	7	B	C	9	A

“Count to Infinity” Probléma

- Módosítások a Distance-Vector routing protokollokban
 - 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, ∞) utat küldi vissza B-nek
- Mindkét módszer csak két csomópontból álló ciklust kerül el

Distance table of A			Distance table of B		
A	cost	next hop	B	cost	next hop
B	2	B	A	2	A
C	3	B	C	∞	-

1X

Distance table of A			Distance table of B		
A	cost	next hop	B	cost	next hop
B	2	B	A	2	A
C	∞	-	C	∞	-

Link State Protokoll

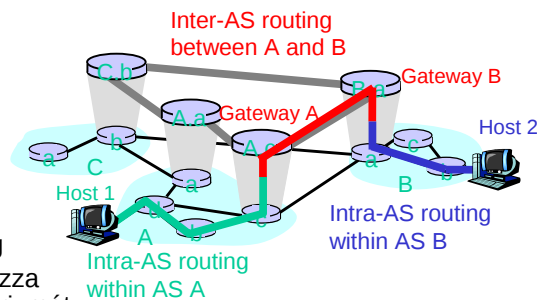
- Minden Link State router
 - tárolja a hálózat topológiáját
 - egy nem-elosztott legrövidebb utak algoritmust használ
- A routerek **Link State Packets (LSP)** által cserélik ki információikat
- LSP tartalmazza
 - az LSP-t létrehozó router IP címét
 - a költségét minden direkt szomszédjához
 - sorozatszámot (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

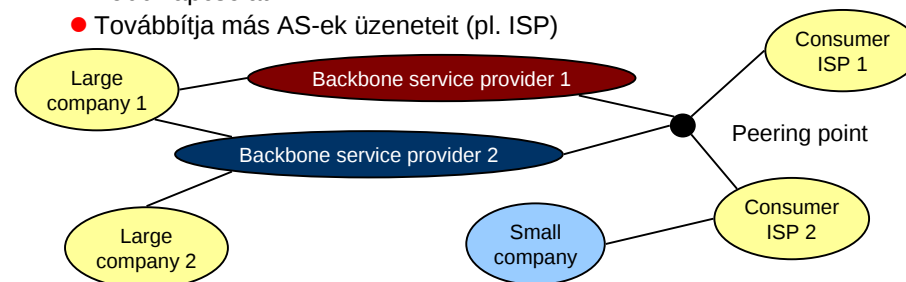
- Nagy hálózatokhoz két hierarchia szint:
 - Lokális terület és gerinchálózat (backbone)
 - Lokális: Link-state advertisement
 - Minden csomópont csak az irányt számítja ki más lokális területek hálózataihoz
- Local Area Border Router:
 - A saját lokális területeik távolságait foglalják össze
 - Ezeket 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

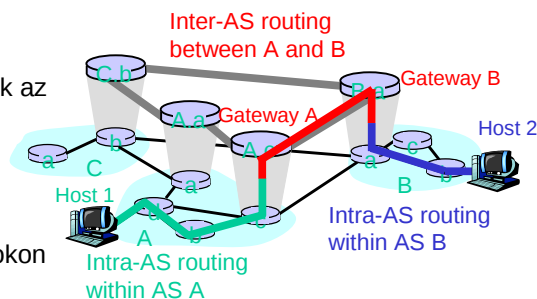
Autonóm rendszerek (AS) típusai

- 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)

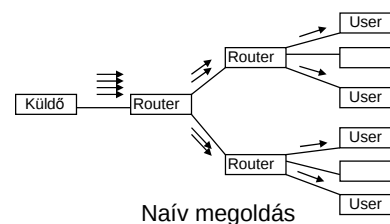
- Az inter-AS routing standard BGPv4
- Path Vector protokoll
 - Hasonló a Distance Vector protokollhoz
 - Minden Border Gateway meghirdeti minden szomszédjának (peers) az egész utat (AS-ek sorozata) a célig (advertisement)
 - TCP-t használ
- Amikor Gateway X az utat Z-hez Peer-Gateway W-nek küldi
 - 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

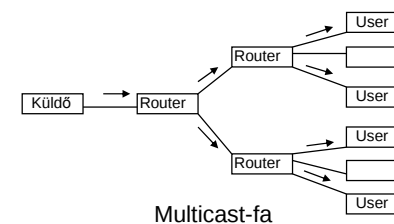
- Broadcast routing
 - Egy csomagot (másolatot) minden más csomópontnak el kell küldeni
 - Megoldások:
 - A hálózat elárasztása (flooding)
 - Jobb: Konstruáljunk egy minimális feszítőfát
- Multicast routing
 - Az adatokat egy küldőtől egyidejűleg több fogadóhoz kell eljuttatni
 - Real time Streaming, Video-On-Demand,
 - Telefon-, Videokonferencia (all-to-all multicast),...
 - IP D címozstály:
 - 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 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 felelőssége segítségével:
 - Minden linken csak egyszer továbbítódik egy csomag.
 - A routerek döntenek el, hogy egy csomagot több linken is továbbítanak-e.



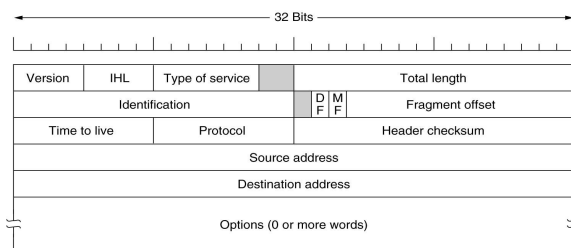
Naív megoldás



Multicast-fa

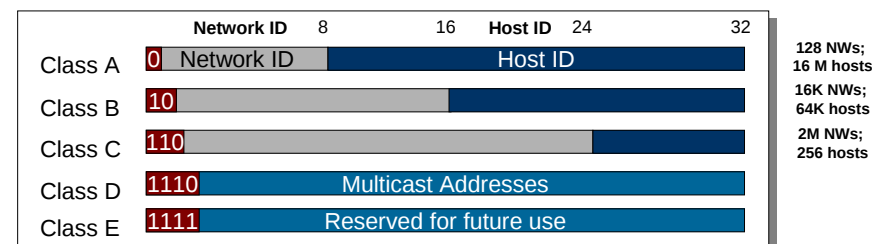
IPv4-Header (RFC 791)

- Version: 4 = IPv4
- IHL: fejléc hossz
 - 32 bites szavakban (>5)
- Type of Service
 - Optimalizálás „delay”, „throughput”, „reliability”, „monetary cost” szerint
- Checksum (csak a fejlécnek)
- Source és destination IP cím
- Protocol: azonosítja a megfelelő protokollt
 - pl. TCP, UDP, ICMP, IGMP
- Time to Live (TTL):
 - maximális hop szám



IPv4 címek

- Osztály alapú címzés (1993-ig)
 - 5 fix osztály, melyek mindegyikét egy prefix azonosítja
 - A, B, C osztály: fix hosszúságú hálózat prefix és host-ID
 - D a multicast osztály; E: lefoglalt

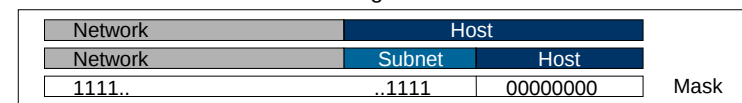


IPv4 címek

- Problémák:
 - A és B osztályú hálózatok sok állomást tartalmazhatnak, ami
 - a routerek számára nehezen kezelhető → **subnetting**
 - Elfogynak a címek: a címosztályok sok címet pazarolnak el
 - pl. egy szervezetnek 1000 állomással már B címosztályra lenne szüksége, ami elpazarol 64K-2K=62K címet → **classless addressing**

IPv4 címek -- Subnetting

- Képzeljünk el egy szervezetet, amely B címosztállyal rendelkezik és a szervezetben belül több LAN van (pl. Egyetem különböző Karai,...)
 - A központi routernek nem kell tudni minden állomásról, csak az egyes LAN-ok egy routeréről → új hierarchiaszint
- Subnetting
 - Bevezetünk alhálózatokat (subnet), melyeket az IP host-részéről leválasztott bitek azonosítanak (a hálózaton kívülről nem látható!)
 - A lokális routereknek tudni kell, hol van ez a leválasztás: ez a **subnet mask** által adható meg



- Pl.: a subnet mask 11111111.11111111.11111111.00000000 azt jelenti, hogy
 - az IP cím 10000100.11100110.10010110.11110011
 - a 10000100.11100110 hálózatban a 10010110 alhálózatban
 - a 11110011 host-ot azonosítja

IPv4 címek -- Classless Inter Domain Routing (CIDR) (RFC 1519)

- A címoszályok nagyon sok IP címet pazarolnak
- 1993 óta: Classless Inter Domain Routing (CIDR)
 - A hálózat cím és a host-ID hossza szabadon adható meg hálózati maszk által.
 - Pl.: hálózati maszk 11111111.11111111.11111111.00000000
 - az IP cím 10000100.11100110.10010110.11110011
 - a 10000100.11100110.10010110 hálózatban
 - a 11110011 host-ot azonosítja
 - Ezek valódi hálózatokat jelentenek amit a többi router is lát – a subnetting-gel ellentétben
 - az IP csomagokba nem kell kiegészítés
 - többi routernek kezelni kell ezeket a változó hosszúságú hálózati címeket (forwarding és a routing algoritmus)

IPv4 címek -- CIDR

- Route aggregation
 - A BGP, RIP v2 és OSPF routing protokollok különböző hálózatokat egy prefix által kezelhetnek
 - Pl. minden hálózat, melynek a prefixe 10010101010* az X szomszédos routeren keresztül érhető el

IP cím lefordítása MAC címre: ARP (RFC 826)

- Address Resolution Protocol (ARP)
- IP cím MAC címre fordítása
 - Broadcast a LAN-ban, lekérdezni azt, hogy melyik állomáshoz tartozik az adott IP cím
 - Csak az a csomópont válaszol, amelyhez az IP tartozik, a MAC címmel
 - A router akkor a csomagot oda ki tudja szállítani

Circuit Switching vagy Packet Switching

- Circuit Switching (vonal kapcsolás)
 - Egy kapcsolat létrehozása felhasználók között kapcsoló állomások által
 - vonalak explicit hozzárendelésével
 - vagy erőforrások explicit hozzárendelésével, pl. idő slots
 - Quality of Service egyszerű, kivéve a
 - kapcsolatfelépítés
 - Probléma
 - statikus hozzárendelés
 - a kommunikációs médium inefficiens kihasználása dinamikus terhelés esetén
 - Felhasználás
 - Telefon, telegráf, mobil telefon

Circuit Switching vagy Packet Switching

- Packet Switching (csomag kapcsolás)
 - Az IP alapelve
 - Az adatokat csomagokra osztja és a csomagokat küldő/fogadó információval ellátva egymástól függetlenül küldi
 - Probléma: Quality of Service
 - A kapcsolat minősége az egyes csomagoktól függ
 - A csomagok puffereződnek vagy el is veszhetnek
 - Előny:
 - A médium hatékony kihasználása dinamikus terhelés esetén
- Összefoglalva
 - Packet Switching gyakorlatilag minden felhasználási területen felváltotta a Circuit Switching-et
 - Ok: a médium hatékony kihasználása