

Az internet ökoszisztémája és evolúciója

Tartalom

- Az útvonalválasztás alapjai
 - elosztott/központi, statikus/dinamikus, link-state/path-vector
 - tartományon belüli/tartományok közötti útválasztás, együttműködés
- Az üzleti modell leképezése AS-útvonalakra
 - valley-free routing: engedélyezett/tiltott utak, a valley-free utak kiszámítása
 - prefer-customer, legrövidebb AS-út, policy routing

Az útvonalválasztás alapjai

Routing versus forwarding

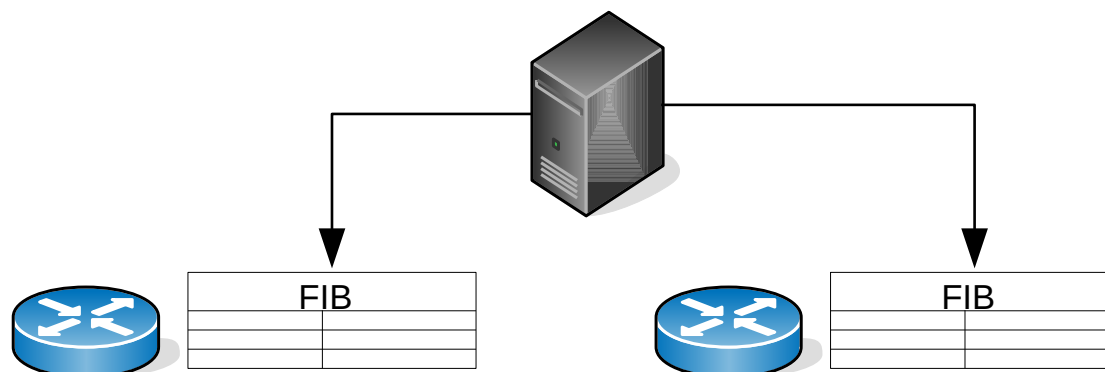
- A routing és a forwarding fogalma gyakran összekeveredik
- **Csomagtovábbítás (forwarding):** egyenként, célcím alapján, a FIB (Forwarding Information Base) szerinti következő IP címre (hop-by-hop)
- Csomagtovábbítás csak helyesen beállított és karbantartott FIB alapján történhet
- **Útvonalválasztás (routing):** topológia-terjesztés, FIB kitöltése/fenntartása/frissítése
- Dedikált routing protokollok végzik

Statikus versus dinamikus

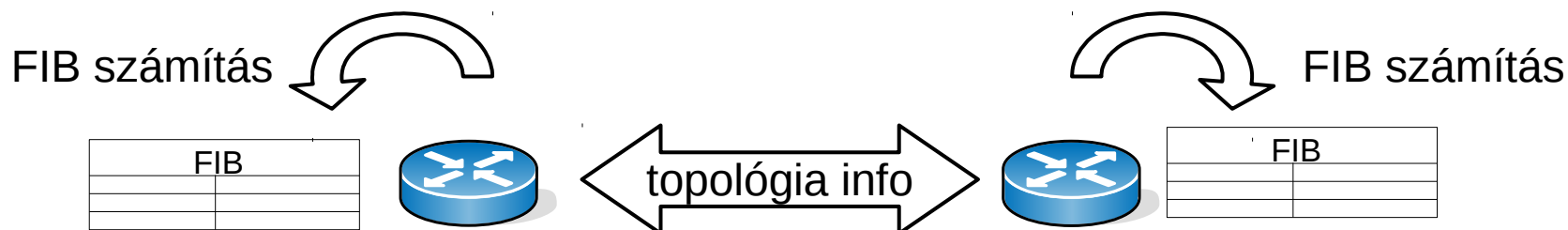
- **Statikus útválasztás:** operátor manuálisan konfigurál (például CLI-n) – **ritkán használt**
 - egyszerű, tetszőleges beállítás eszközölhető
 - de nem skálázódik és nem adaptálódik a topológiaváltozásokhoz (hiba esetén)
- **Dinamikus útválasztás:** dedikált routing protokoll állítja a FIBeket – **elterjedten használt**
 - adaptív és skálázható
 - de az utakat megszabja a protokoll (pl. csak per-destination legrövidebb út)

Dinamikus: központi vs elosztott

- **Központi útválasztás:** dedikált szerver (route-szerver, PCE, SDN controller) állítja a FIBeket



- **Elosztott útválasztás:** a routerek topológia-leírókat cserélnek és egyenként állítanak FIBet



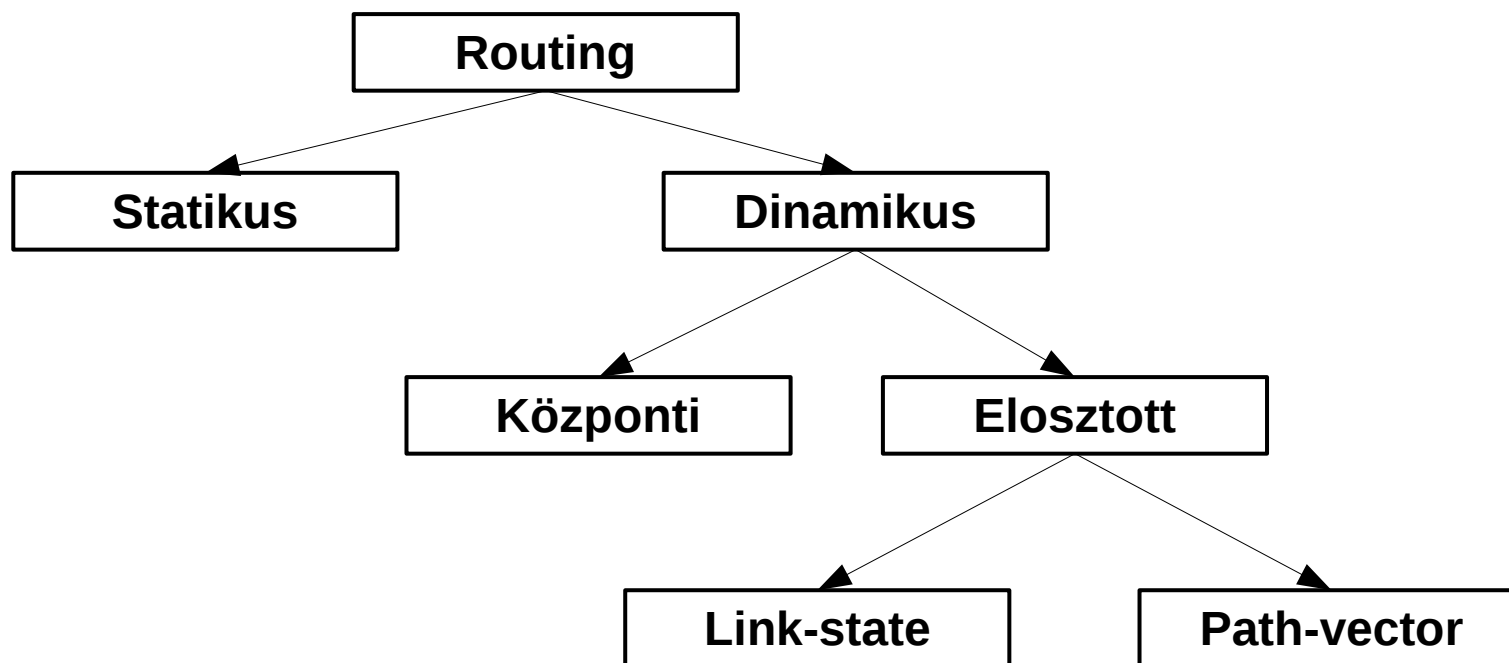
Elosztott útválasztás menete

- Szomszédos routerek felfedezik egymást
- A routerek megosztják egymás közt a topológiával vagy a lehetséges útvonalakkal kapcsolatos információikat (routing state)
- A hálózatileíró információkat a szomszédos routerek folyamatosan frissítik (így egy elosztott routing state adatbázist fenntartva)
- Minden router minden általa ismert prefixre kiválasztja a legjobb útvonalat
- A legjobb útvonal next-hop-ját letölti a FIBbe

Elosztott: link-state vs path-vector

- **Link-state:** hálózatileíró = topológia
 - minden router tudja a gráfot, rugalmas útválasztás valamilyen „távolság” szerint
 - de csak ha minden router megegyezik a policy-ben! (pl. legrövidebb út az ASen belül)
- **Path-vector:** hálózatileíró = átjárható útvonalak
 - next-hop: az ismert utak közül a „legjobb”
 - nem kell „távolság” metrika
 - routerenként egyedi policy állítható be
- **Distance-vector:** kihalóban

Routing protokollok



ASen belül és ASek között

- Az internet hierarchiaszintjein különböző routing protokollok futnak
- **Intra-domain routing:** ASen belüli hosztok/routerek közti utak kiépítése
 - **Interior-Gateway Protocol (IGP)**
 - tipikusan ASen belül a routing policy homogén (pl. legrövidebb út): **link-state**
- **Inter-domain routing:** ASek közti útvonalak
 - **Exterior Gateway Protocol (EGP)**
 - heterogén útválasztási szempontok: **path-vector**

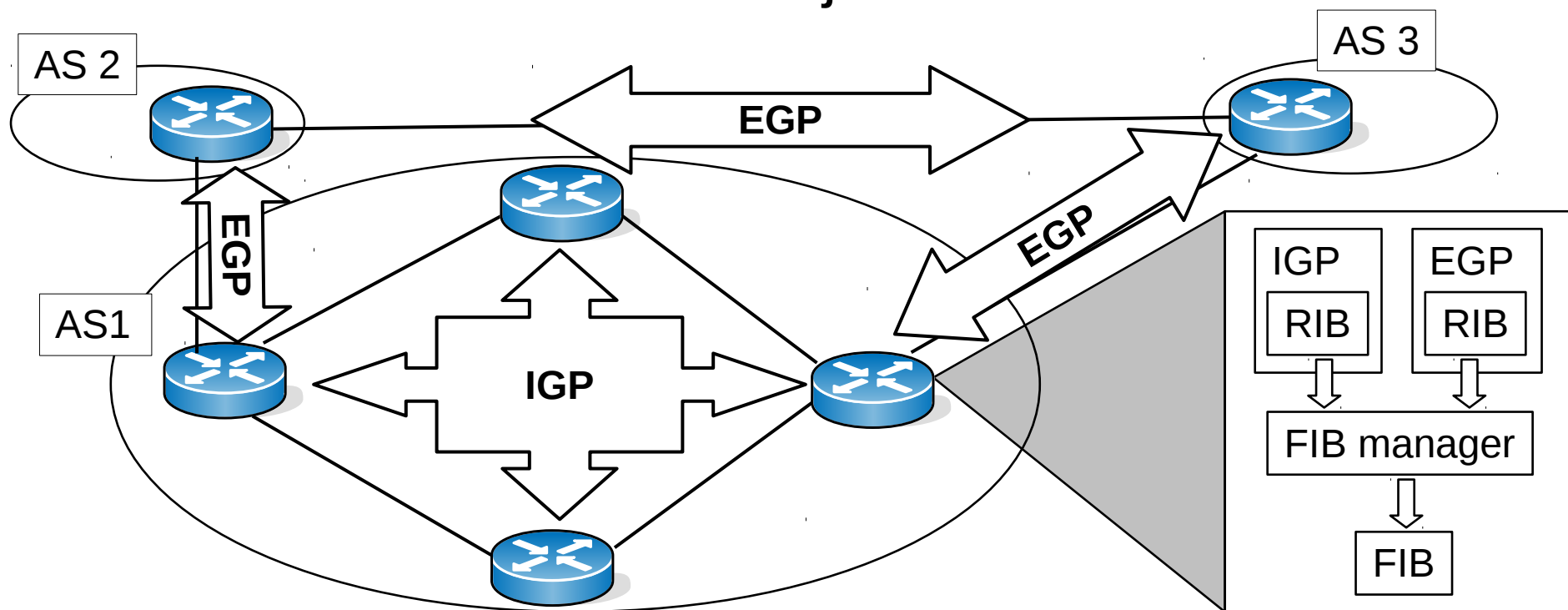
A „Nagy négyes”

	Link-state	Vector
IGP	OSPF, IS-IS	RIP (kb.)
EGP		BGP

- OSPF: Open-Shortest Path First
- IS-IS: Intermediate-System-to-Intermediate-System
- RIP: Routing Information Protocol
- BGP: Border Gateway Protocol

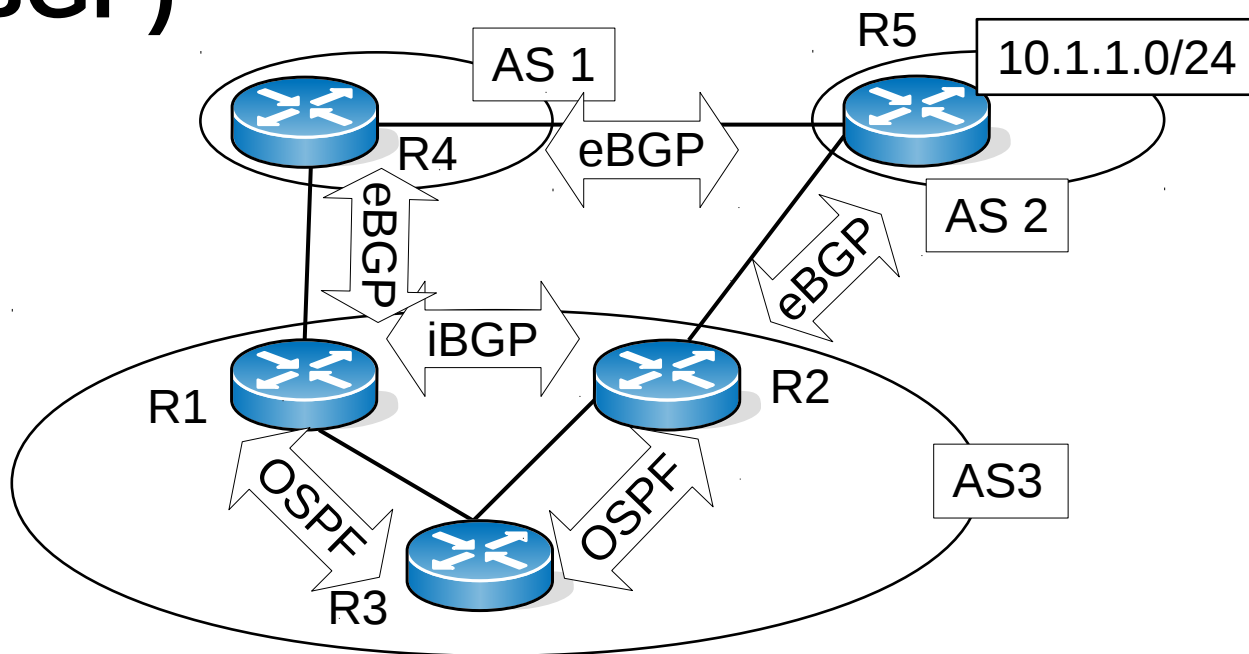
IGP/EGP együttműködés

- Több routing protokoll fut párhuzamosan
 - legalább egy EGP és egy IGP
 - mindegyik saját RIBet épít
 - a router ebből szintetizálja a FIBet



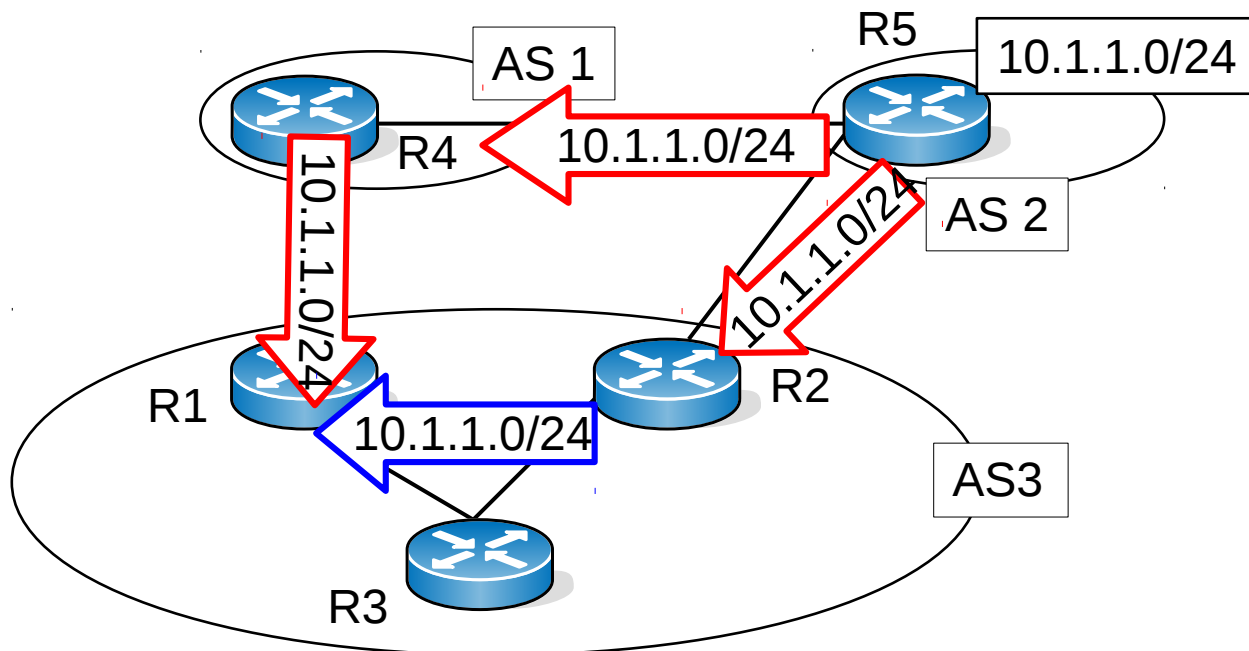
IGP/EGP együttműködés

- Alább R1 két módon érhet el egy külső hosztot
 - AS1-en keresztül (R4 felé)
 - vagy AS2-n keresztül (R2 felé majd R5-ön át)
- A külső útvonalakat ASen belül is terjeszteni kell:
iBGP (Internal BGP)
- Az ASek közt futó BGP-t ezért gyakran **eBGP**-nek (External BGP) hívják



IGP/EGP együttműködés

- Például AS2 hirdeti a 10.1.1.0/24 prefixet
- R5 eBGP-n elküldi a hirdetést R2-nek és R4-nek, amely szintén eBGP-n elküldi R1-nek
- R2 a hirdetést iBGP-n továbbítja R1-nek (iBGP-n, hiszen ASen belüli routerek közt történik)

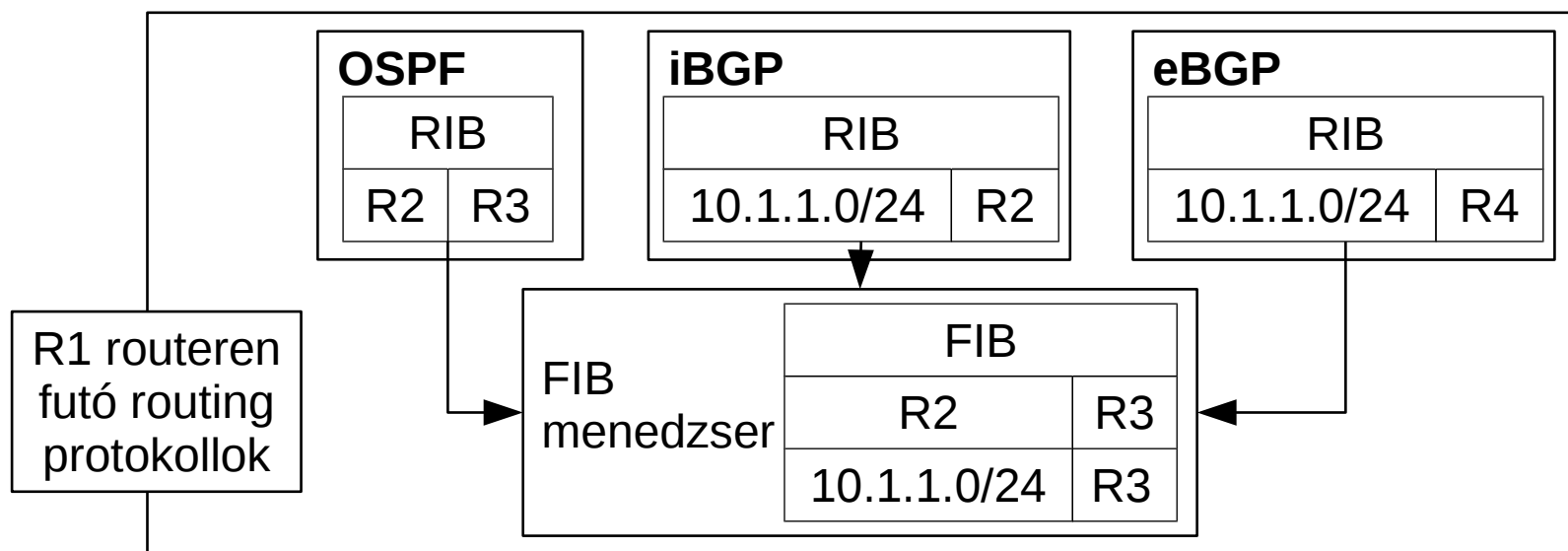
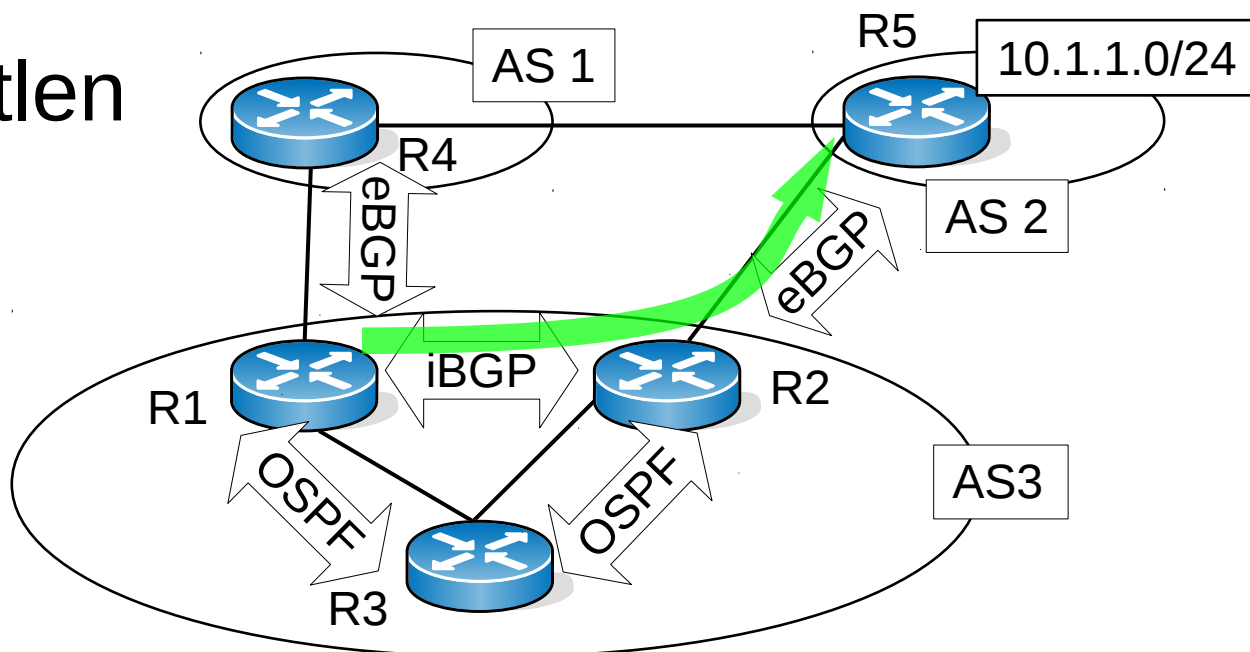


IGP/EGP együttműködés

- R1 a `10.1.1.0/24` prefixre több hirdetést kap
 - R4-től kapott hirdetés az eBGP RIB-ben, R2-től kapott hirdetés az iBGP RIB-ben (sokszor nem különülnek el ezek, Ciscón például közös az eBGP-iBGP RIB)
 - de mivel R2 közvetlenül nem szomszéd, ezért az OSPF RIB-ben meg kell keresni $R1 \rightarrow R2$ next-hopot
- Az operátor dönti el, melyik utat preferálja \rightarrow FIB
- A „távoli” next-hopokat rekurzívan visszafejti
 - `10.1.1.0/24` prefixhez R2 a next-hop (iBGP), R2-höz pedig R3(OSPF), így `10.1.1.0/24`-höz végül az R3 next-hopot rendeljük a FIBben

IGP/EGP együttműködés

- Ha AS3 a közvetlen AS2 felé vezető utat preferálja
- Az iBGP hirdetést fogadja el



IGP/EGP együttműködés

- A **FIB menedzser** feladata az egyes routing protokollok RIBjeiből összeállítani a FIBet
- Az operátor súlyokat rendel a protokollokhoz
- Ha egy prefix több RIBben is megtalálható, a **kisebb súlyú** protokoll nyer

lokális < statikus < OSPF < eBGP < iBGP

- A lokális útvonalakat mindig azonnal elfogadjuk
- eBGP < iBGP miatt a szomszéd AS-tól kapott utat preferáljuk az AS-beli routertől kapott út előtt: **hot-potato routing** (később)

Az AS-szintű üzleti modell leképezése útvonalakra

Inter-domain routing

- **Útválasztási preferencia (routing policy):** egy ISP AS-AS szintű üzleti stratégiájának leképezése útválasztási szabályokra
- A bonyolult útválasztási preferenciák kifejezéséhez hatékony útválasztási protokoll szükséges → a BGP

Hogyan lehet a tranzit-peer kapcsolatokat illetve a „prefer-customer” szabályt leképezni az útválasztás nyelvére?

AS kapcsolatok: tranzit vs peer

- Két AS tipikusan vagy tranzit vagy peer kapcsolatot hoz létre
 - **tranzit:** globális internet-hozzáférés pénzért
 - **peer:** „ingyen” adatcsere a két szolgáltató és azok összes előfizetője között
- Most csak ezekkel a tipikus esetekkel fogunk foglalkozni

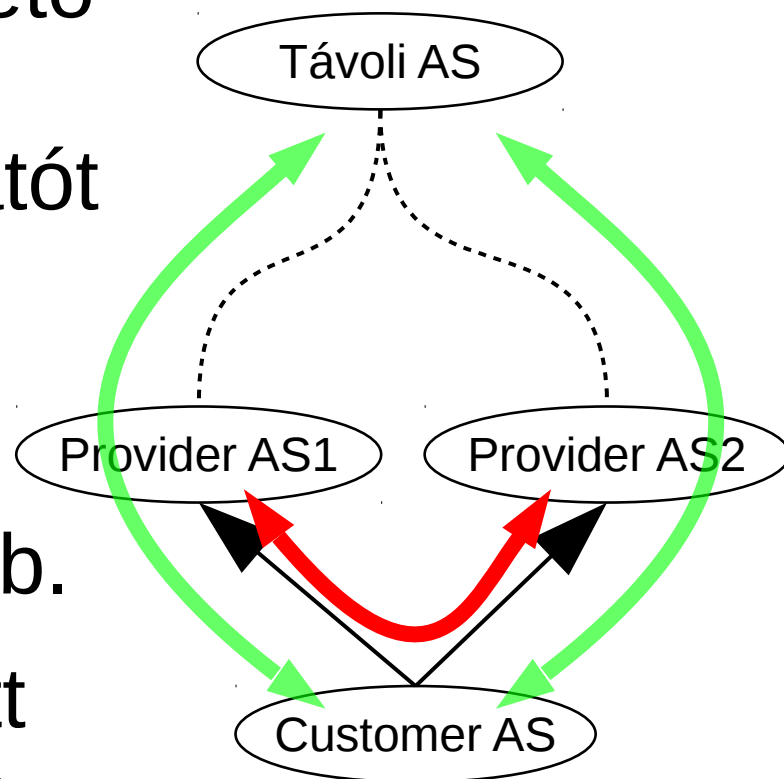
Milyen AS-szintű útvonalakat eredményeznek a fenti üzleti megfontolások?

Engedélyezett és tiltott utak

- A megvalósuló utak leképezik az ISP-k gazdasági/üzleti érdekeit
- **Engedélyezett út:** útvonal egy AS-en keresztül, amely megegyezik annak gazdasági érdekeivel
 - pl. bármilyen tranzit út egy szolgáltató AS-en keresztül → profit
- **Tiltott út:** a gazdasági érdekekkel ellentétes út
 - pl. két tranzit szolgáltató közt, előfizetőn át
 - az előfizető nem szerződött ilyen forgalomra → anyagi veszteség

Tranzit: megengedett/tiltott utak

- **Engedélyezett út:** az előfizető és bármely AS között (beleértve magát a szolgáltatót illetve annak bármely előfizetőjét is): *Customer AS ↔ Provider AS1*, *Customer AS ↔ Távoli AS*, stb.
- **Tiltott út:** szolgáltatók között az előfizető AS-en keresztül: *Provider AS1 → Customer AS → Provider AS2*

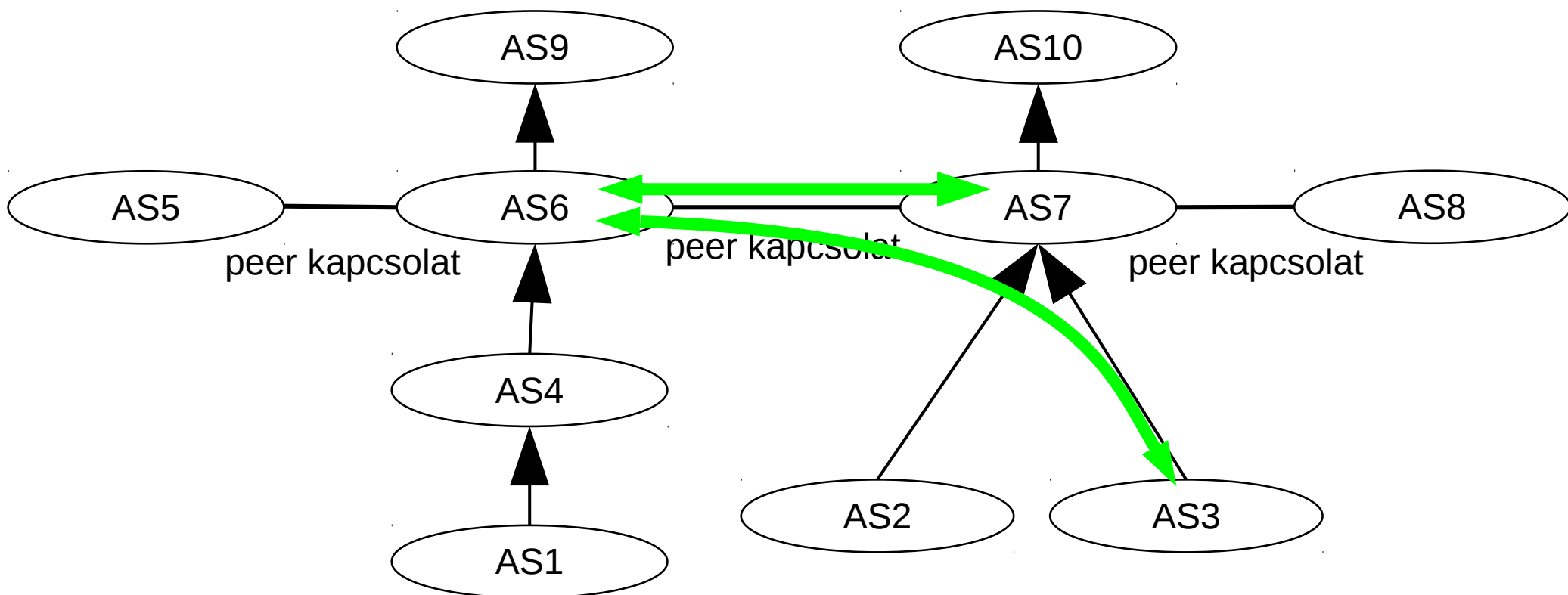


Peer kapcsolat: engedélyezett utak

1) Két peer AS között: AS6-AS7, AS7-AS8, ...

2) Előfizető és a peer között: AS4-AS7, AS6-AS3

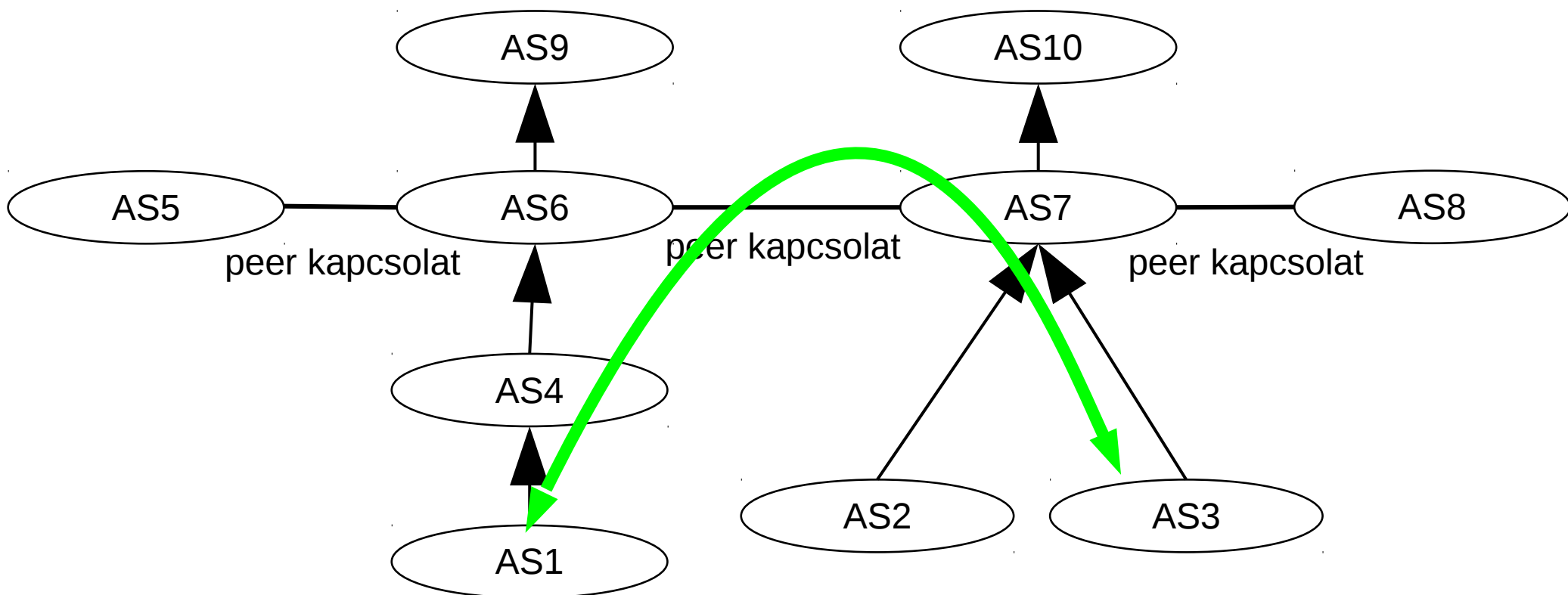
- előfizető előfizetője is OK: AS1-AS7



Peer kapcsolat: engedélyezett utak

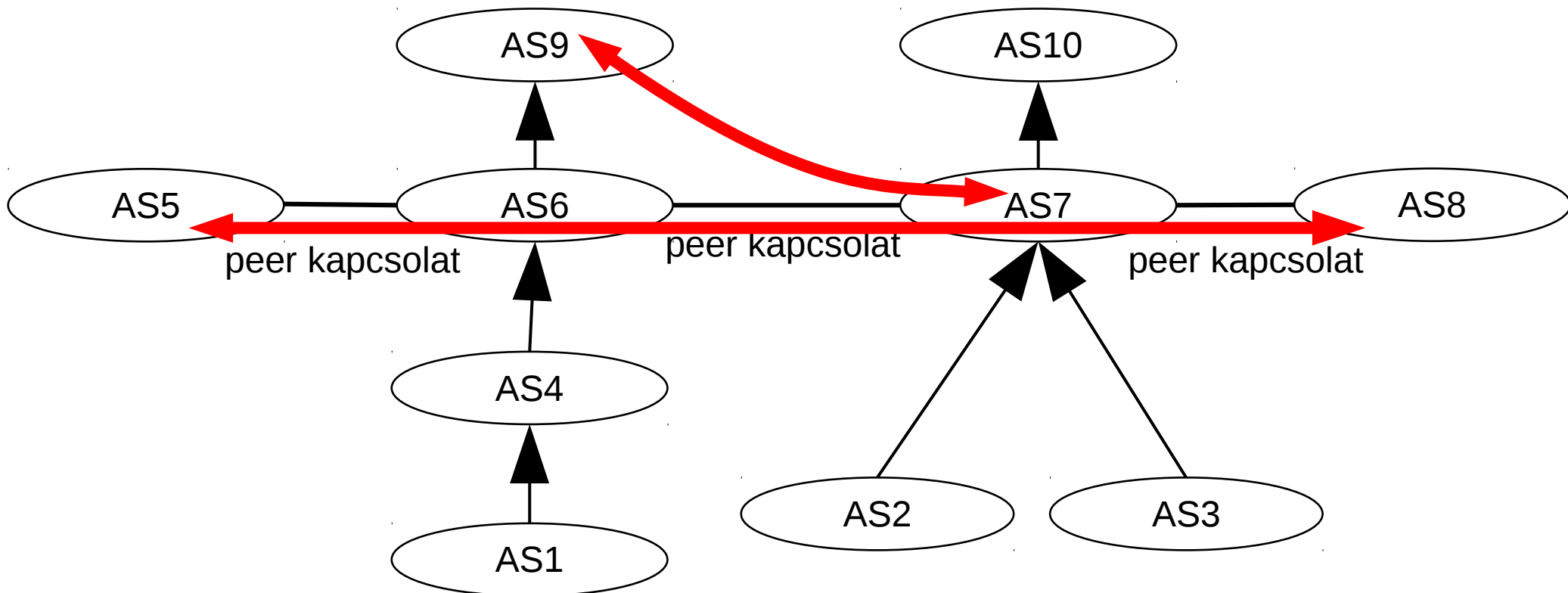
3) Peer AS-ek bármely két előfizetője között:
AS4-AS3

- tranzitív viszony: AS1-AS2, AS1-AS3 is OK



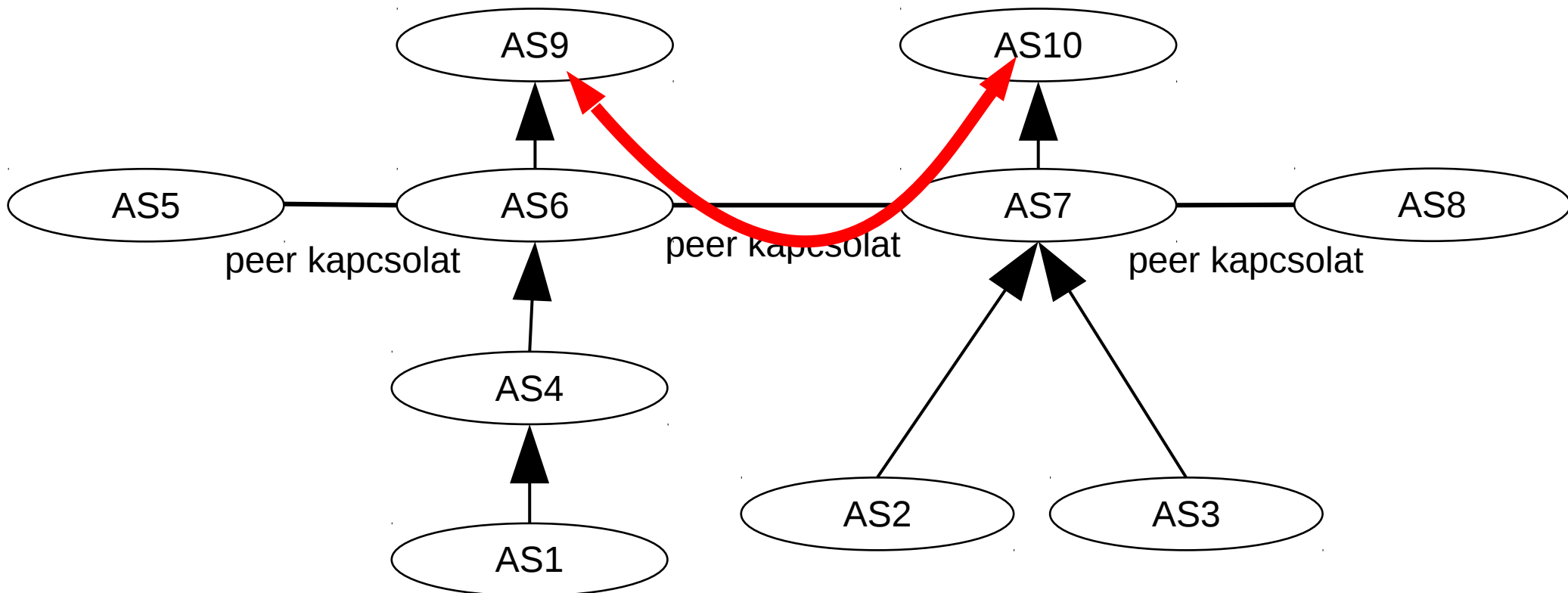
Peer kapcsolat: tiltott utak

- 1) Nem közvetlen peer kapcsolatban levő AS-ek között: AS5-AS7, AS5-AS8
- 2) AS szolgáltatója és peer-e között: AS9-AS7



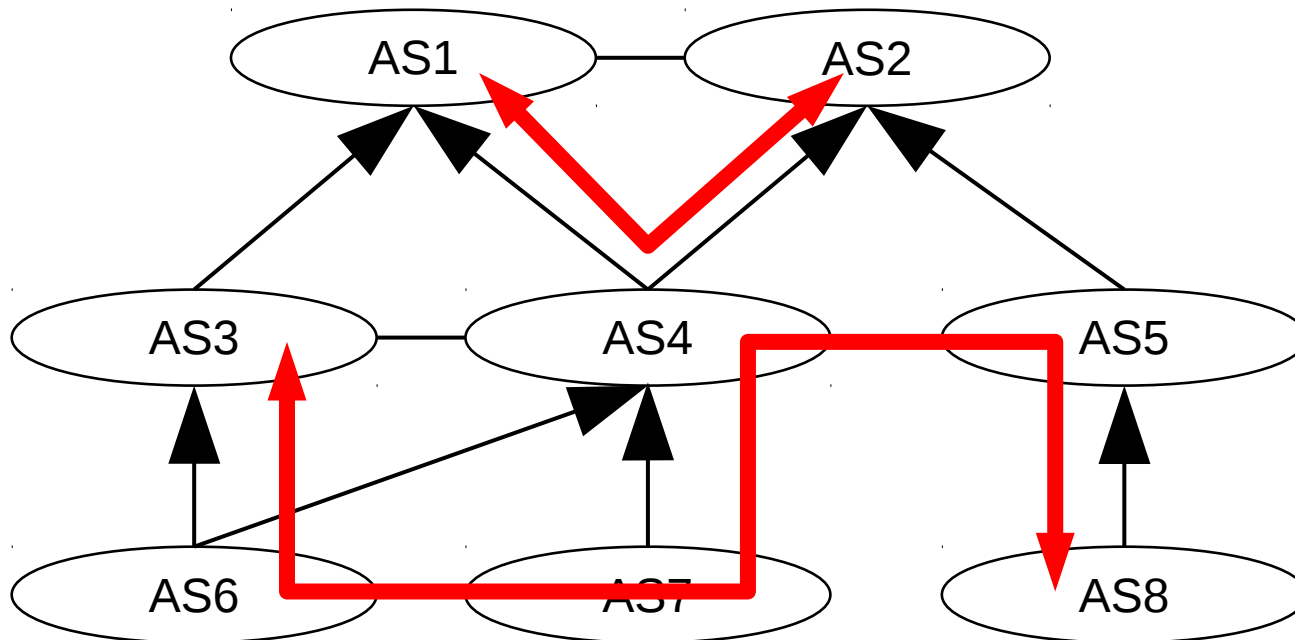
Peer kapcsolat: tiltott utak

3) AS bármely szolgáltatója és a peer AS bármely szolgáltatója között (AS9-AS10), mert ez tranzit szolgáltatás lenne, amiért pénz járna



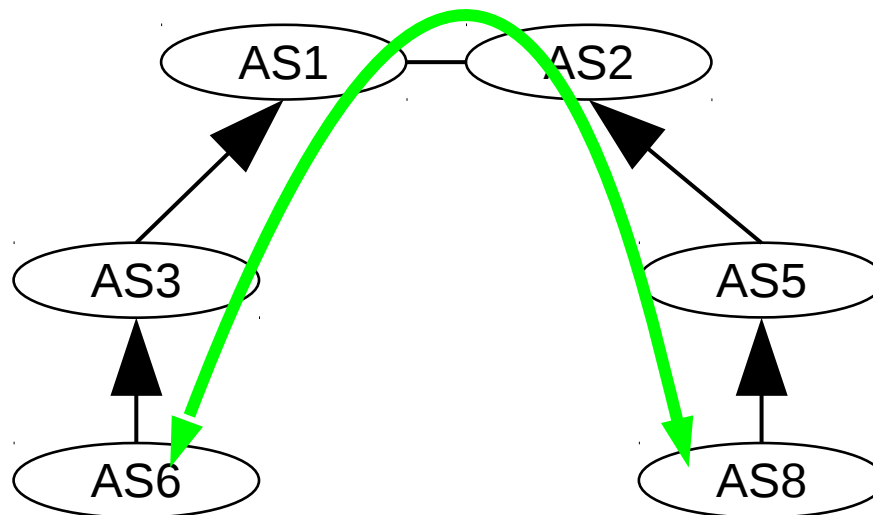
Valley-free routing

- A fenti kritériumok szerinti útvonalválasztás együttes neve: **valley-free routing**
- Az elnevezés háttere: a valley-free routing tiltja a „völgyeket” a tranzit hierarchiában



Valley-free routing: szabályok

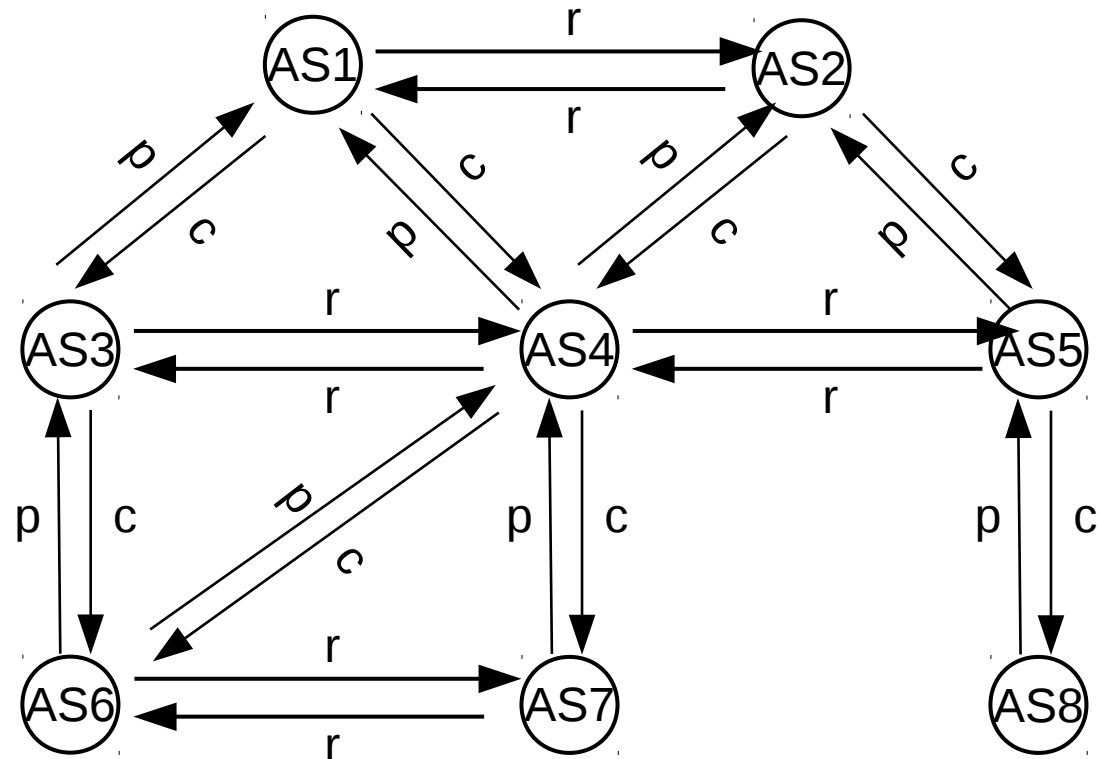
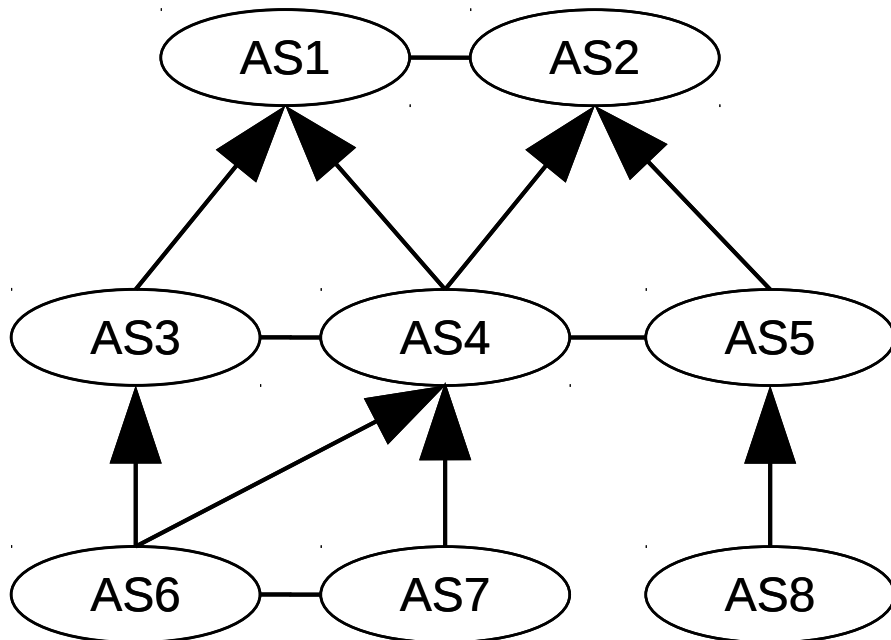
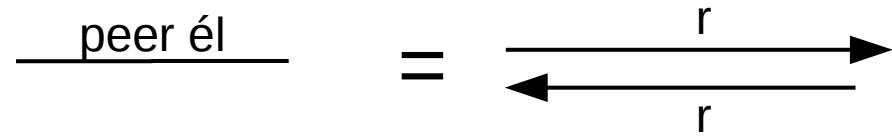
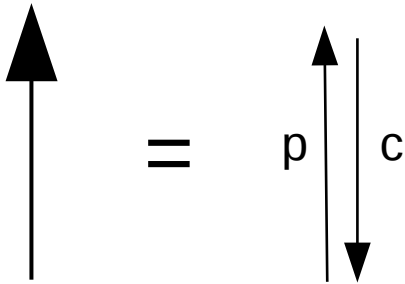
- **Def.:** egy útvonal megengedett, ha sorrendben
 - először tartalmaz nulla vagy több *előfizető* → *szolgáltató* élt
 - majd nulla vagy egy *peer* ↔ *peer* élt
 - végül bármennyi *szolgáltató* → *előfizető* élt



Valley-free routing reprezentáció

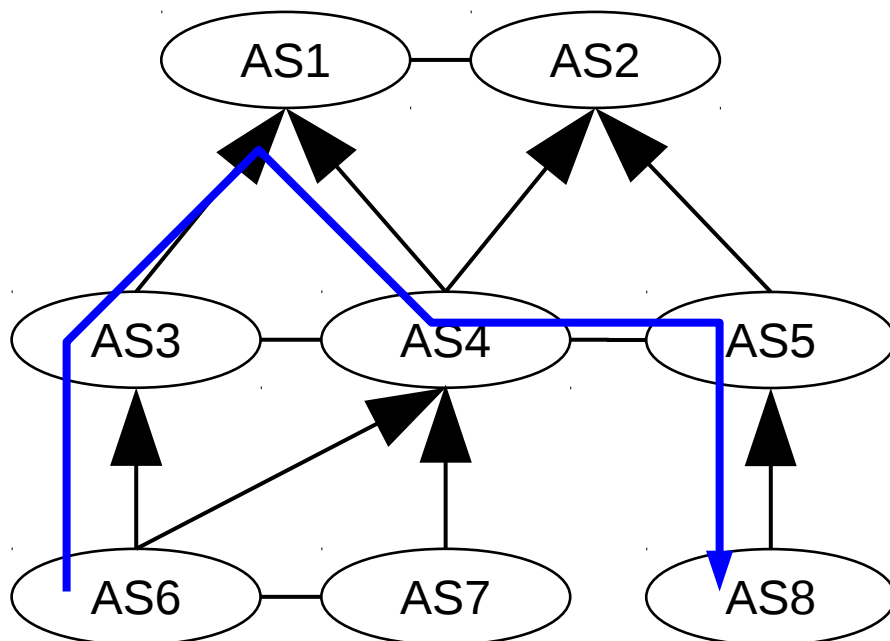
- Hogyan kell a valley-free útvonalakat számolni?
- Építsünk egyszerű címkézett AS-szintű **gráf-reprezentációt**
- A gráf pontjai az ASek, élei pedig az alábbiak
 - a tranzit él az *előfizető* → *szolgáltató* irányban: irányított él p (**provider**) címkével
 - a *szolgáltató* → *előfizető* irányban: irányított él c címkével (**customer**)
 - a *peer* él bármely irányban: irányított él r címkével (**peer**)

Valley-free routing reprezentáció



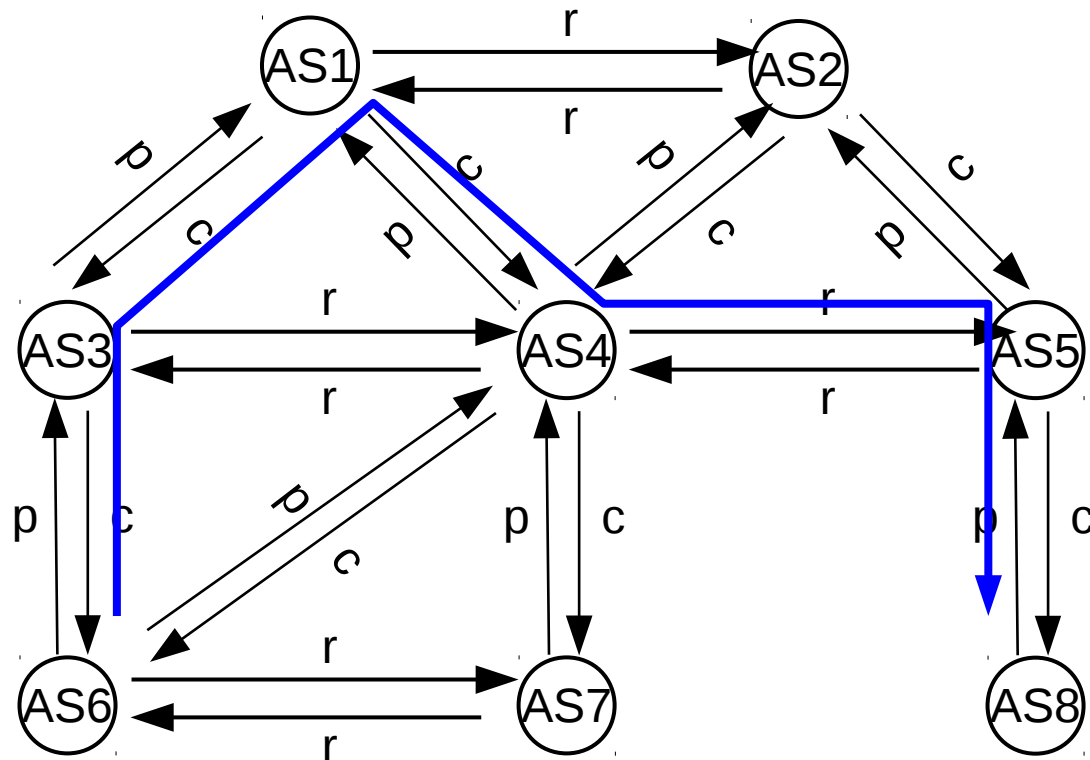
Valley-free routing számítása

- A gráfrepresentáció egyszerűen használható annak eldöntésére, hogy egy útvonal megengedett-e a valley-free routingban
- Megengedett például az alábbi útvonal?



Valley-free routing számítása

- Írjuk fel a gráfrepresentációt az útvonallal



- Írjuk fel az irányított útvonal éleinek címkéjét:

p, p, c, r, c

Valley-free routing számítása

- Zárójelezzünk hátulról és végezzük el az alábbi műveletet

+	<i>c</i>	<i>p</i>	<i>r</i>	∞
<i>c</i>	<i>c</i>	∞	∞	∞
<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	∞
<i>r</i>	<i>r</i>	∞	∞	∞

$$p+(p+(c+(r+c))) = p+(p+(c+r)) = p+(p+\infty) = p+\infty=\infty$$

- Tiltott ha az eredmény ∞ , megengedett egyébként (így az eredmény: az út tiltott!)

Valley-free routing számítása

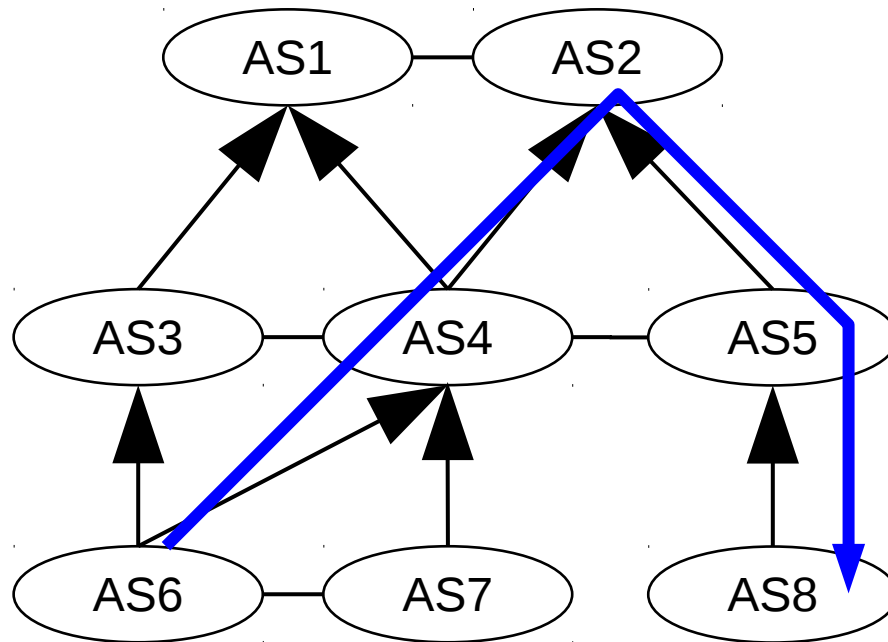
+	<i>c</i>	<i>p</i>	<i>r</i>	∞
<i>c</i>	<i>c</i>	∞	∞	∞
<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>	∞
<i>r</i>	<i>r</i>	∞	∞	∞

- Tiltott útvonal: ha cp , cr , rp , vagy rr részútvonalat tartalmaz
- **Tétel:** egy valley-free útvonal megfelel az alábbi reguláris kifejezésnek:

$$p^*r^*c^*$$

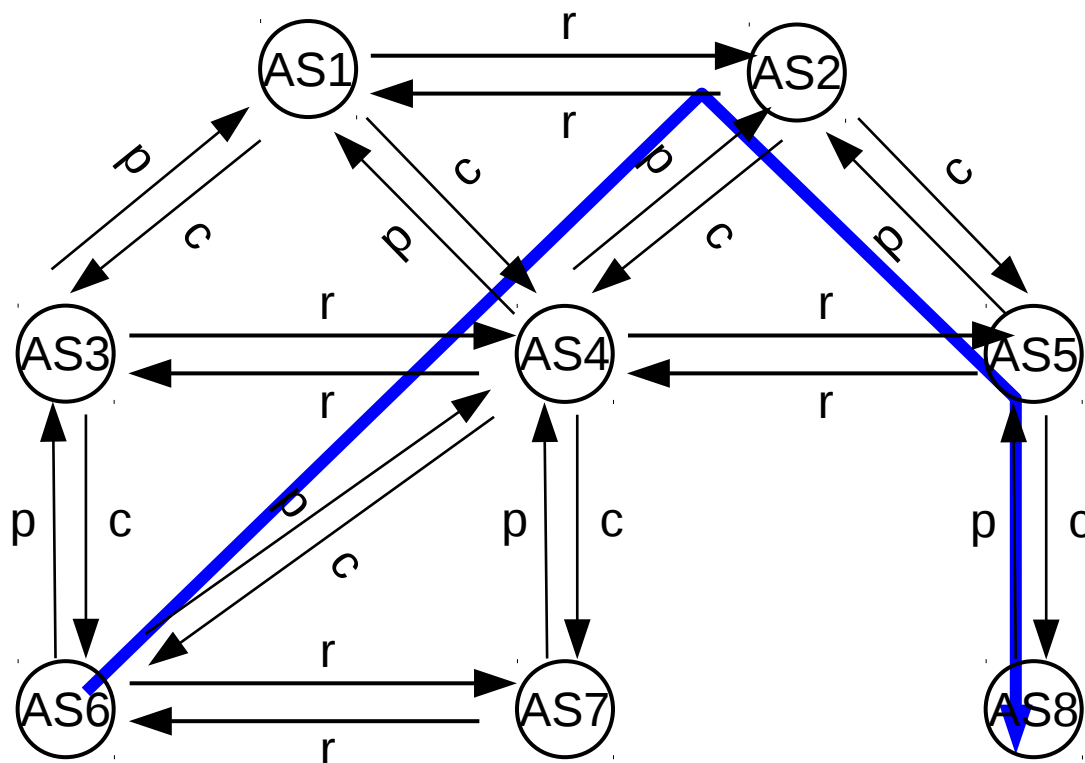
Valley-free routing: példa

- Megengedett-e az alábbi $AS6 \rightarrow AS8$ útvonal?



Valley-free routing: példa

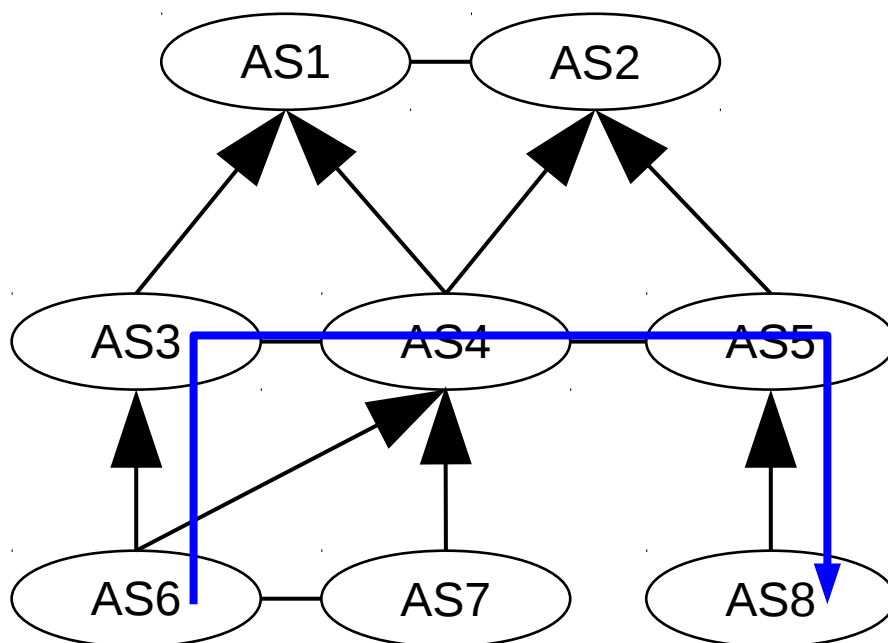
- A gráfrepresentációból kiolvasható az útvonal címkéje: $p+(p+(c+c))=p+(p+c)=p+p=p$
- Az útvonal megengedett



Valley-free routing: példa

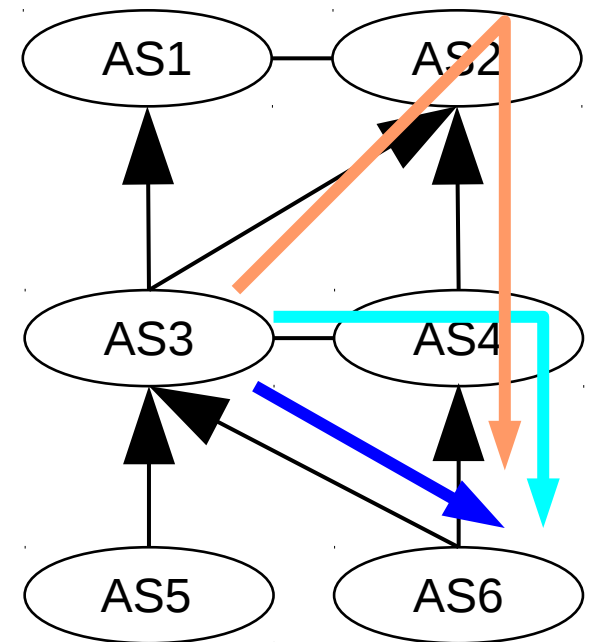
- De az alábbi $AS6 \rightarrow AS8$ útvonal már nem megengedett

$$p+(r+(r+c))=p+(r+r)=p+\infty=\infty$$



Útvonalak preferenciája

- A példában AS3-nak 3 valley-free útvonala van AS6 felé:
 - AS2 tranzit szolgáltatón keresztül
 - AS4 peer kapcsolaton keresztül
 - vagy a közvetlen AS6 linken keresztül
- Hogyan válasszon ezek közül?

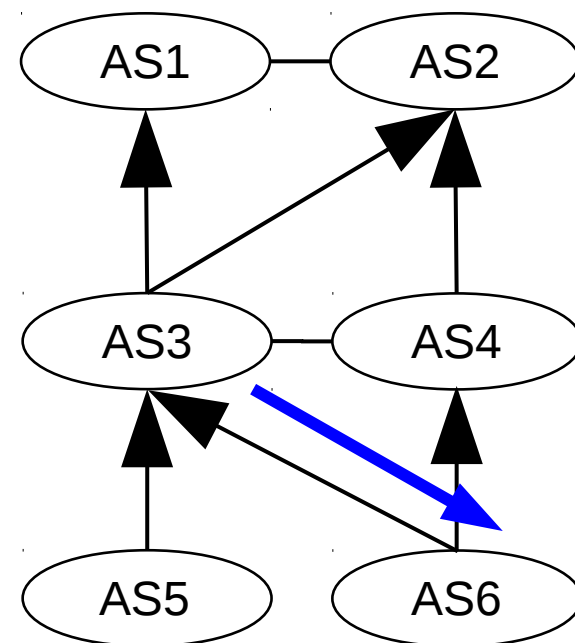


Útvonalak preferenciája

- Ha a szolgáltató felé irányítjuk a forgalmat: tranzit díjat kell fizetni
- Ha egy peer AS felé: a forgalom valószínűleg ingyen van (kivéve paid peering esetén)
 - de ha „túlterheljük” a peer kapcsolatot
 - felborulhat a „szimmetrikus forgalmi igények” kritérium a peering policy-ben
- Ha előfizető felé: biztos, hogy ingyen lesz a forgalom

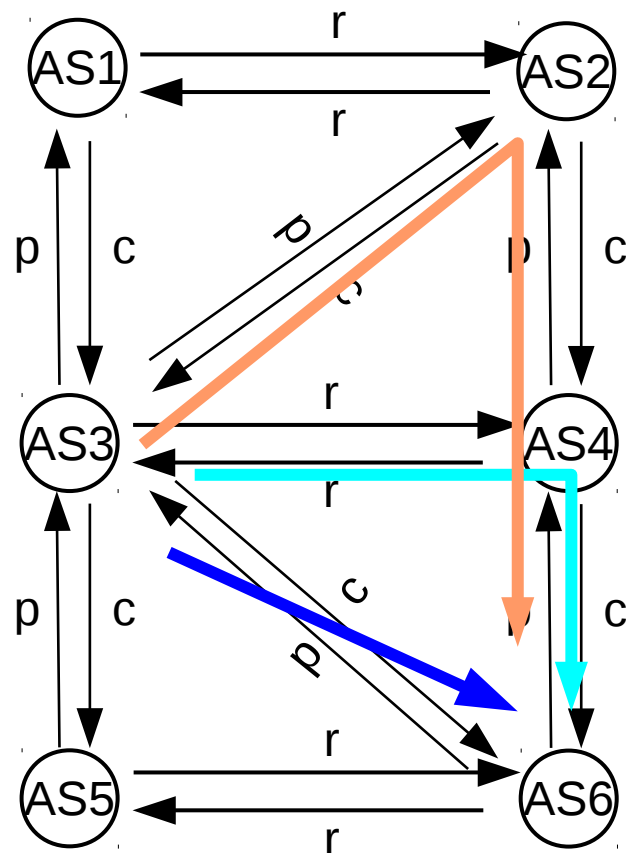
„Prefer customer” szabály

- Tipikusan az ISP-k preferálják az előfizetőn keresztüli útvonalakat
- A preferencia sorrendje:
 1. *szolgáltató* → *előfizető* élel kezdődő útvonal
 2. *peer* → *peer* élel kezdődő útvonal
 3. *előfizető* → *szolgáltató* élel kezdődő útvonal



„Prefer customer” szabály

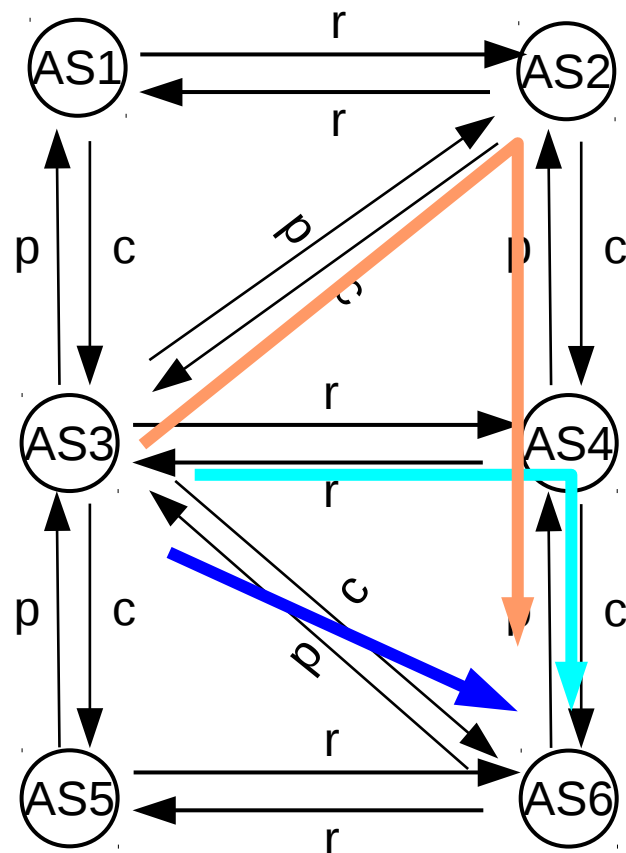
- A gráfmodellben:
 - a szolgáltatón keresztül vezető útvonal:
$$p+(c+c) = p+c=p$$
 - a peer útvonal
$$r+c=r$$
 - a közvetlen út: c
- **Észrevétel:** a kalkulus pont az út első élének címkéjét adja végeredményül



„Prefer customer” szabály

- Legyen $l(P)$ egy P útvonal címkéje, amelyet úgy kapunk, hogy kiértékeljük a + műveletet az út éleinek címkéjén
 - P_p (provider út): $l(P_p) = p$
 - P_c (customer út): $l(P_c) = c$
 - P_r (peer út): $l(P_r) = r$
- **Tétel:** a „prefer customer” leírható az alábbi relációval:

$$P_c < P_r < P_p$$



„Prefer customer” szabály

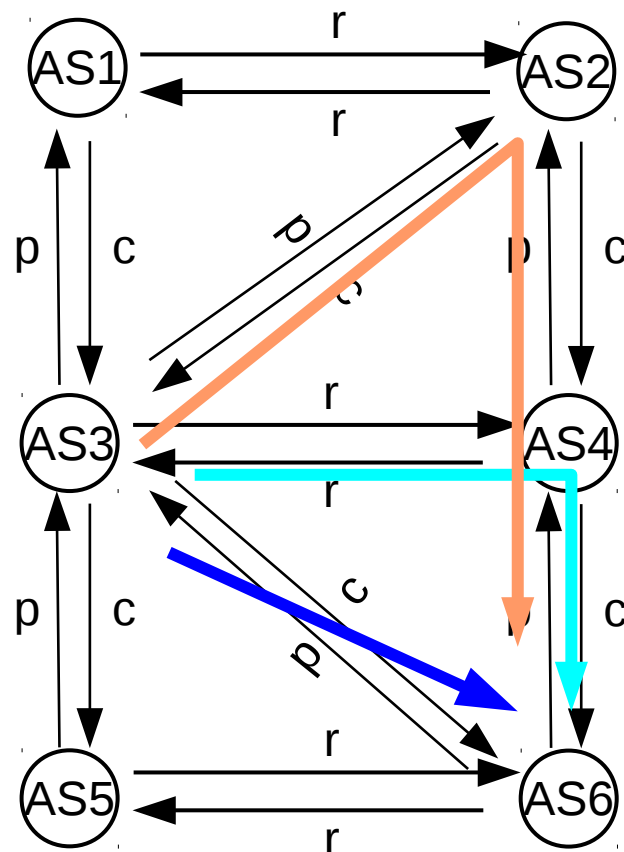
- A példánkban:

$c <$ (előfizető út)

$r+c=r <$ (peer út)

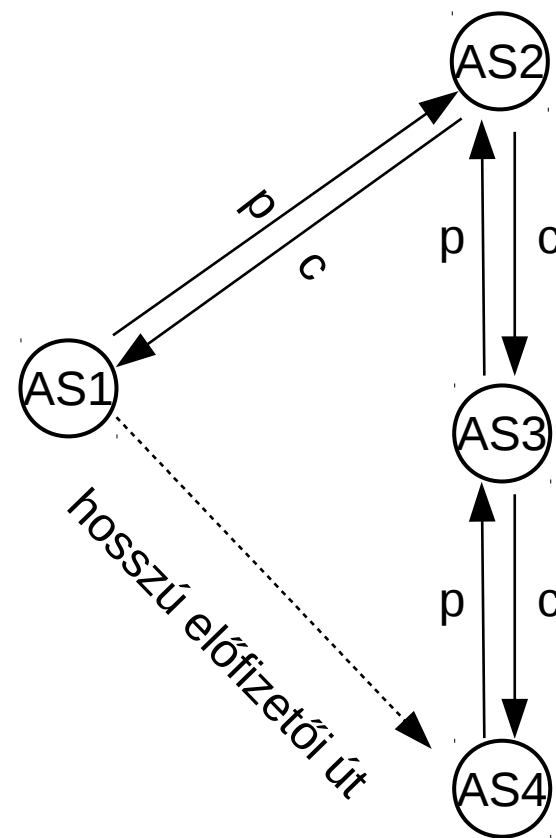
$p+(c+c)=p+c=p$ (tranzit út)

- A relációs jel a kevésbé preferált útvonal felé mutat



„Prefer customer” szabály

- A „prefer customer” szabály nem mindig eredményez „jó” utat
- Ha AS1 és AS4 között előfizetői linkek hosszú sorozata
- Ésszerűbb lehet a rövidebb tranzitot választani AS2-n át
- Az operátor autonóm döntése
- Az ISP policy-k nagyon összetettek lehetnek!



Legrövidebb AS-útvonal

- Ha több, azonos preferenciájú útvonal is elérhető
 - például itt AS1 választhat egy 2 AS-ből álló és egy „nagyon hosszú” előfizetői út közül
- Célszerű az azonos preferenciájú utak közül a rövidebbet választani:
legrövidebb AS útvonal
- De először a preferencia számít!

