

Az internet ökoszisztémája és evolúciója

Tartalom

- Internet szolgáltatási modell
 - AS-szintű üzleti kapcsolatok
 - tranzit szolgáltatás: definíció, jelölés
 - single-homed/multi-homed AS, PA/PI címezés, stub/tranzit AS, tranzit hurkok elkerülése, AS-ek mérete, tranzit árazás
 - peer szolgáltatás: definíció, jelölés, peering policy-k, IXP-k
- Az internet AS-szintű struktúrája

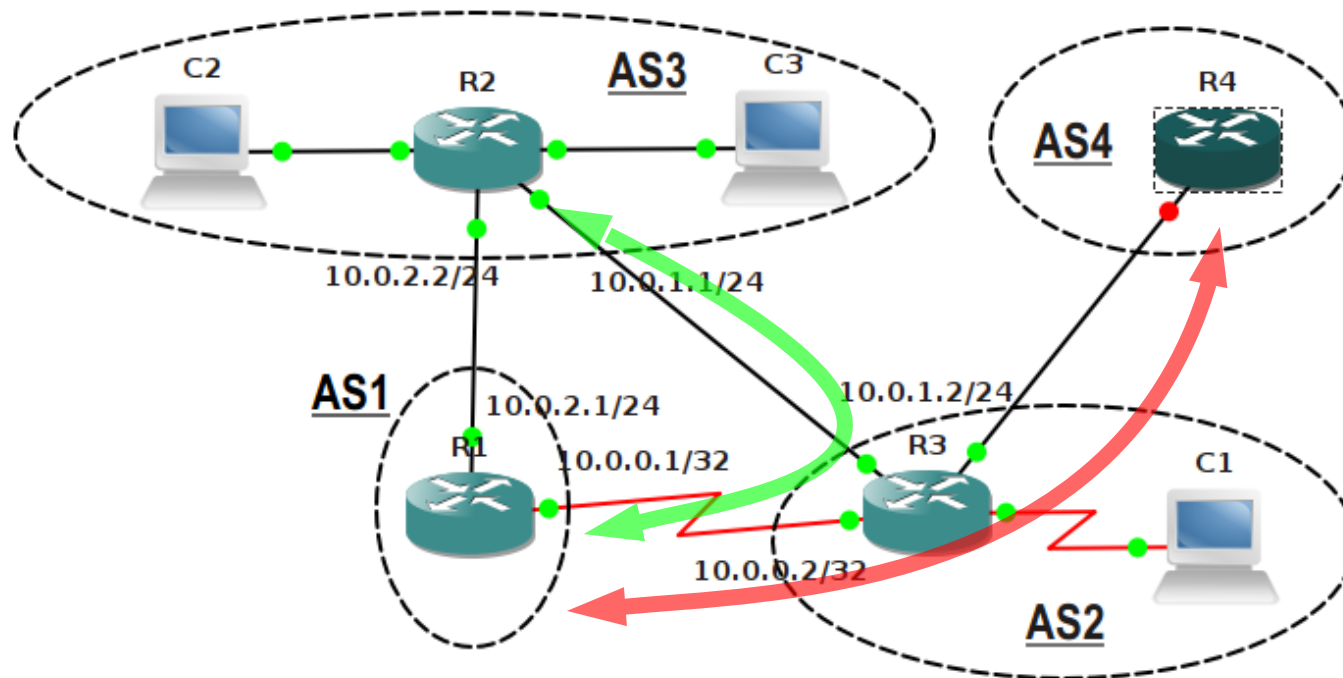
Internet szolgáltatási modell

AS-ek összekapcsolása

- Az internet-szolgáltatás jövedelmező üzlet egy nagyon kompetitív piacon
- A szolgáltatók a hálózataikat **üzleti érdekeik szerint kapcsolják össze**, és ugyanezen üzleti logika szerint **alakítják ki az útvonalakat**
- **Útválasztási stratégia (policy routing)**: üzleti, politikai, biztonsági, megbízhatósági, stb. érdekek érvényesítése az útvonalválasztásban
- Az interneten a forgalom nem a legrövidebb útvonalakon halad (**path inflation**)!

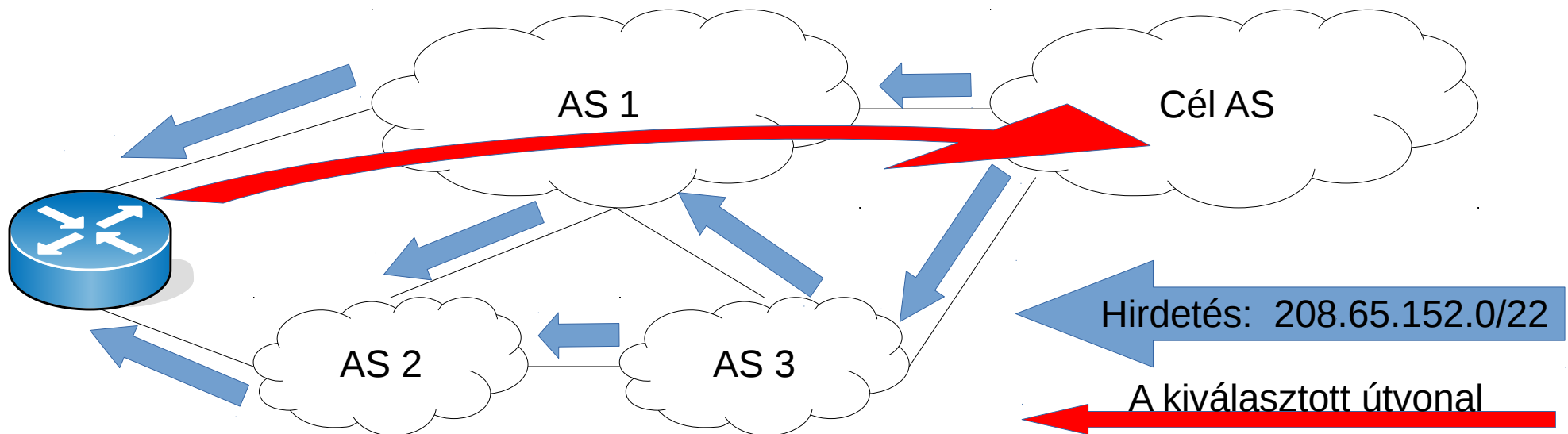
Útválasztási stratégiák

- Valamely AS melyik másik ASek forgalmát engedi át a hálózatán: szolgáltatási modell
- Példa: AS2 továbbítja az AS1–AS3 forgalmat, de blokkolja AS1–AS4 átmenő forgalmát



Tartományok közti útválasztás

- A szolgáltatók IP címtartományt igényelnek
- A tartományt „meghirdetik”, az útválasztó protokoll (BGP) pedig biztosítja, hogy a hirdetés a hálózat minden részébe eljusson
- A routerek útvonalat választanak



A tranzit szolgáltatás

- Két Autonóm Rendszer (AS) összekapcsolásának legegyszerűbb módja
- Egyik AS a **szolgáltató (provider)**, másik AS az **előfizető (customer)**
- A szolgáltató szerződésben vállalja, hogy
 - az előfizető AS forgalmát eljuttatja az internetre csatlakoztatott bármely hoszthoz
 - és az internetről az előfizető AS-hez érkező összes forgalmat eljuttatja hozzá
- Az előfizető cserében forgalomfüggő díjat fizet

A tranzit szolgáltatás

Internet Transit is the business relationship whereby an Internet Service Provider provides (usually sells) access to the global Internet.

From a high-level perspective, Internet Transit can be thought of as a pipe in the wall that says "Internet this way". Customers connect their networks to their Transit Provider, and the Transit Provider does the rest.

(DrPeering)

A tranzit szolgáltatás: címzés

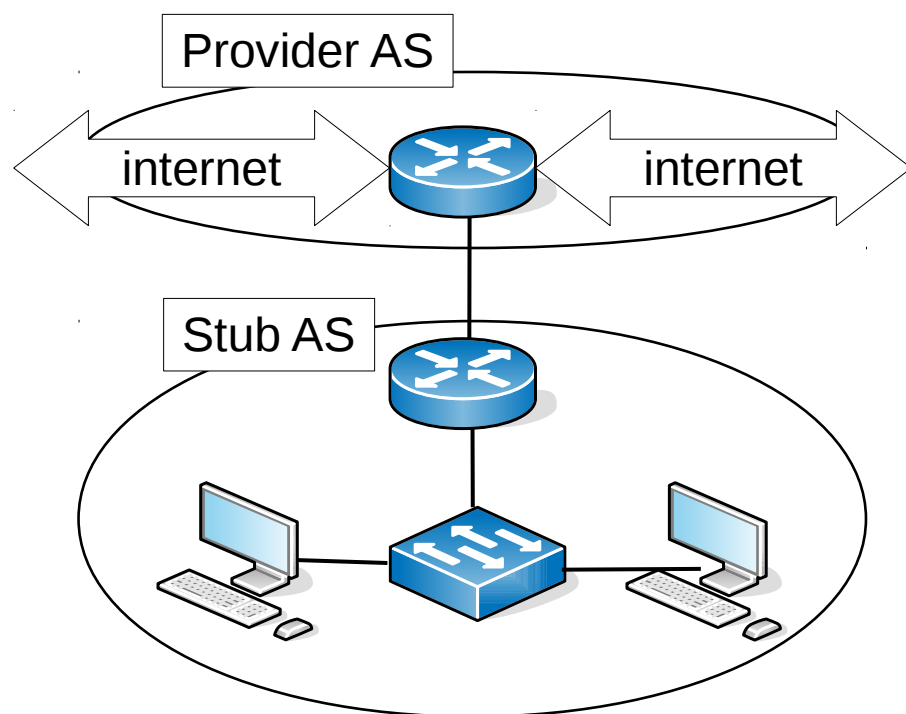
- Az előfizető számára a szolgáltató IP címtartományt ad(hat)
- A tranzit szolgáltató feladata ezután, hogy
 - az előfizető IP címtartományára vonatkozó BGP hirdetést terjessze az interneten
 - a külső címtartományokra vonatkozó BGP hirdetéseket átadja az előfizetőnek
 - majd pedig folyamatosan továbbítsa a „előfizető ↔ internet” forgalmat

Tranzit: a single-homed AS-ek

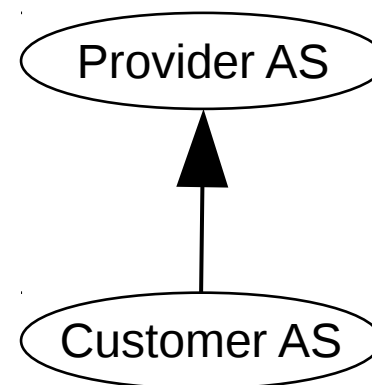
- A legegyszerűbb módja egy AS-t az internethez csatlakoztatni: egyetlen szolgáltatótól tranzit szolgáltatást igénybe venni
- **Single-homed (SH) AS:** olyan AS amely egyetlen AS tranzit szolgáltatását igénybe véve kapcsolódik az internetre

SH AS: sematikus jelölés

- A tranzit kapcsolatokat az előfizetőtől a szolgáltató felé mutató nyíllal jelöljük
- A nyíl a cash-flow (\$) áramlásának irányát adja

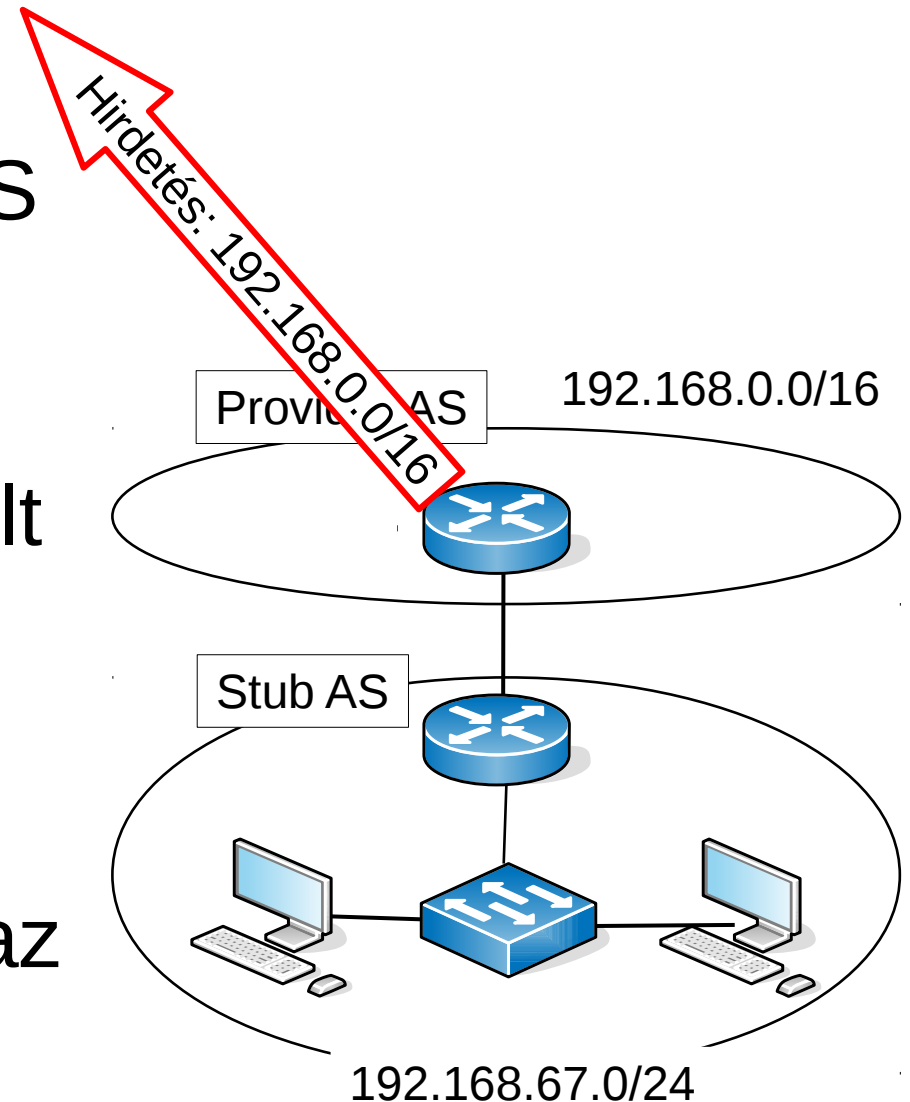


Sematikus jelölés



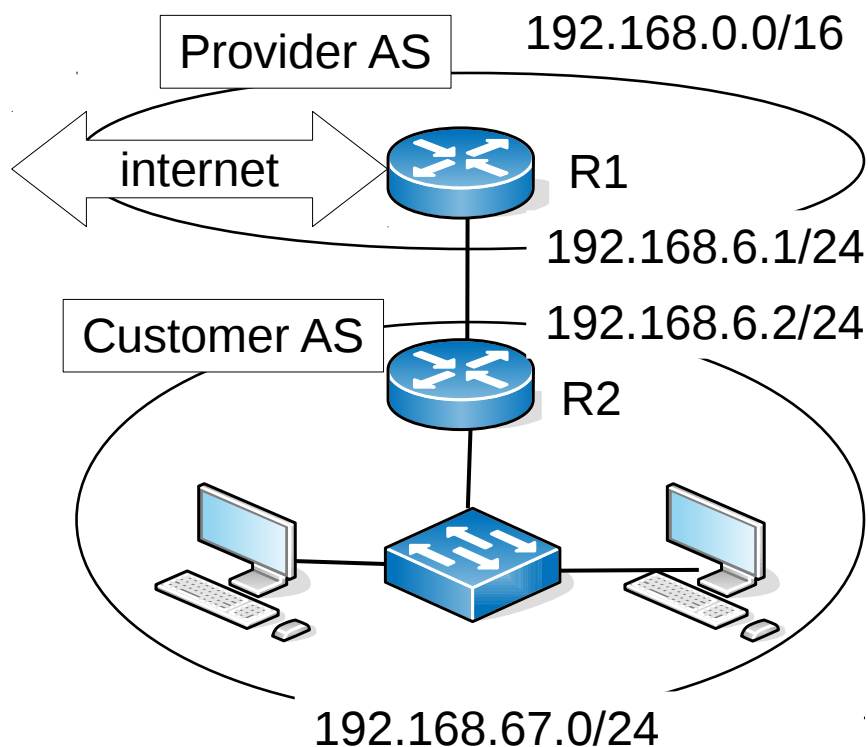
SH tranzit: címzés és routing

- **Címzés:** a SH AS tipikusan a szolgáltató AS címtartományának egy alhálózati prefix-ét kapja
- **Routing:** a SH AS default gateway-e a szolgáltató
- A szolgáltató nem hirdeti külön az SH előfizető címtartományát, hiszen az része a saját címtartományának



SH tranzit: példa

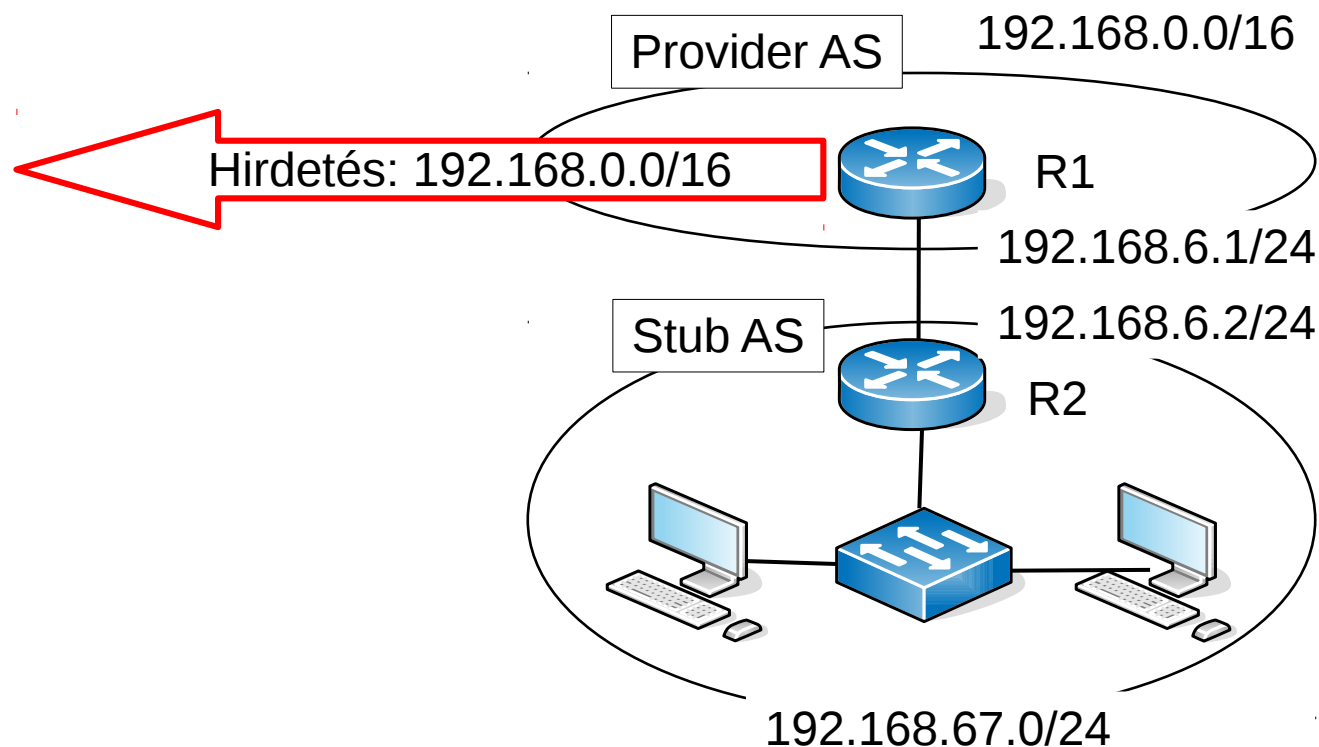
- Tfh. a $192.168.0.0/16$ (egyébként privát) Class C prefix a Provider AS címtartománya
- Így 256 darab $/24$ címtartományt tud kiosztani



- Tegyük fel, hogy a Customer AS a $192.168.67.0/24$ címtartományt kapja
- Határrouterek címei:
R1: $192.168.6.1$
R2: $192.168.6.2$

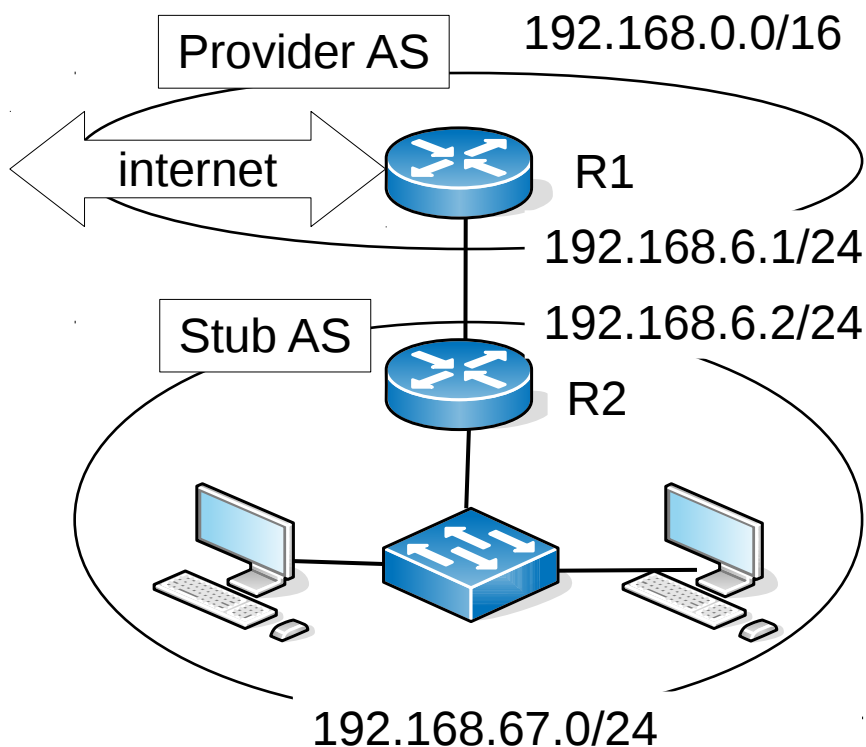
SH tranzit: példa

- A szolgáltató nem hirdeti az előfizető címtartományát külön, hiszen az része a saját címtartományának



SH tranzit: példa

- A szolgáltató az előfizető bejövő forgalmát egy specifikus (**more specific**) FIB bejegyzéssel az előfizető felé irányítja

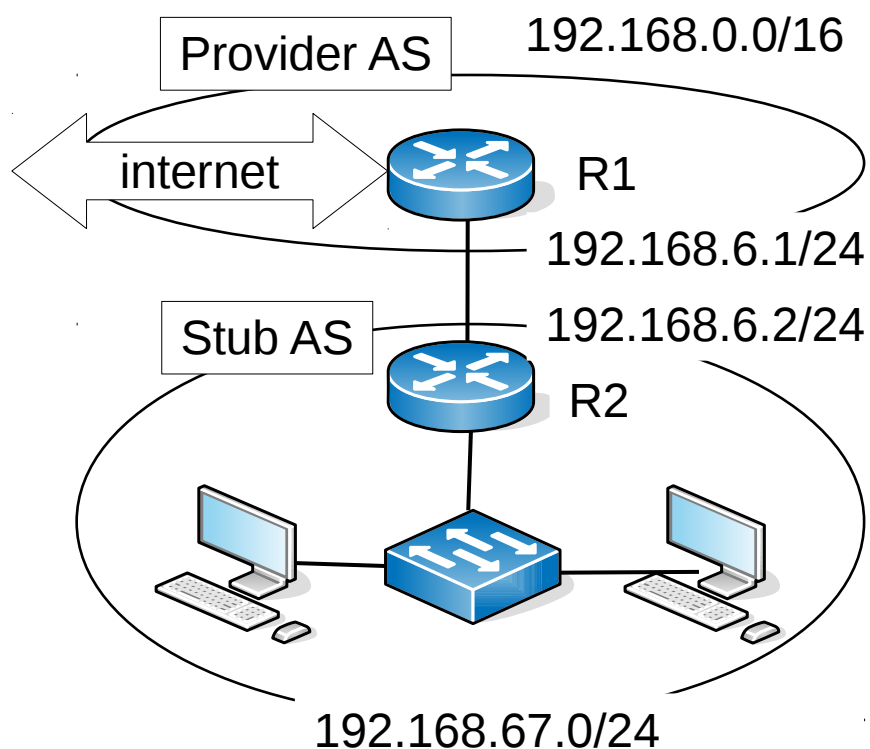


- R1 forgalomtovábbítási táblája:

Prefix	Next-hop
...	...
192.168.67.0/24	192.168.6.2
...	...

SH tranzit: példa

- Az előfizető **default gateway**-t állít be a szolgáltató felé: a default gateway-nél minden bejegyzés specifikusabb = jobban preferált



- R2 forgalomtovábbítási táblája:

Prefix	Next-hop
0.0.0.0/0	192.168.6.1

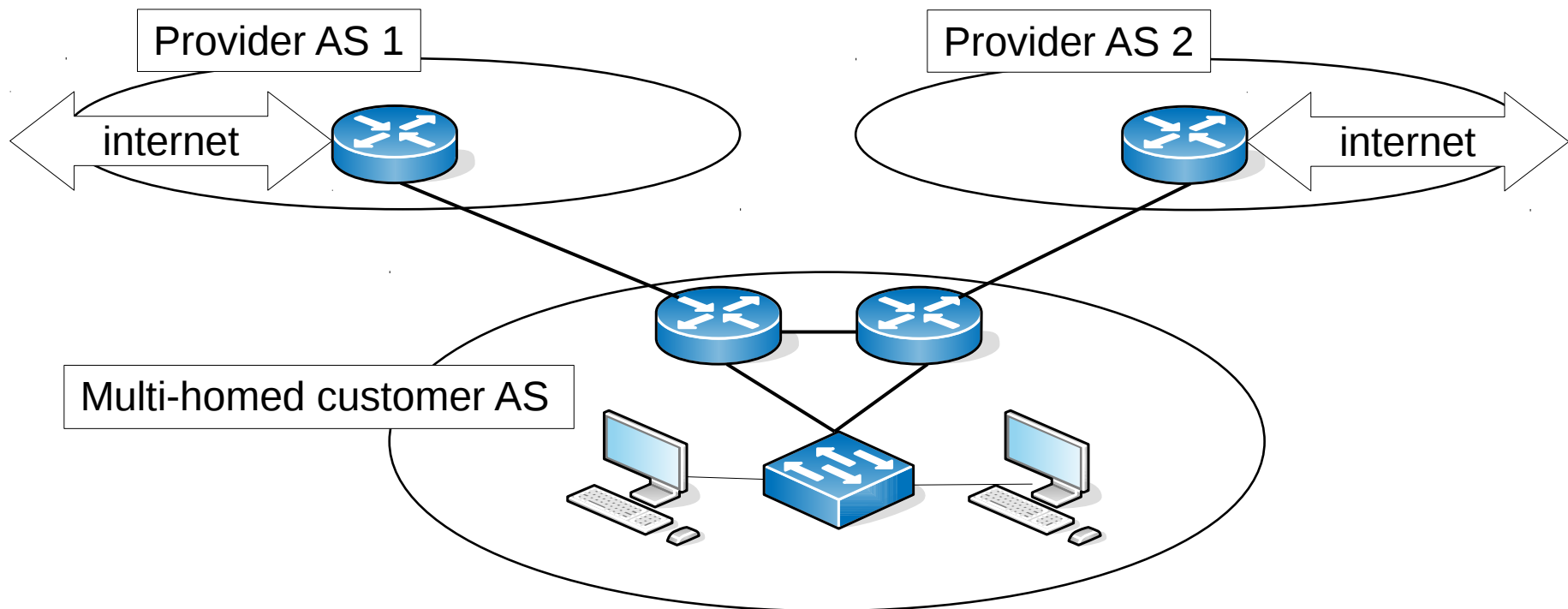
- Ehhez routing protokoll sem kell a SH AS-ben!

Single-homing: értékelés

- Előnyök
 - minimális konfiguráció a SH AS-ben
 - nem keletkezik újabb FIB bejegyzés az internet többi részén
- Hátrányok
 - rossz hibatűrés (ha az access link kiesik, nincs Internet kapcsolat a SH AS-ben)
 - esetleg drága

Multi-homed AS-ek

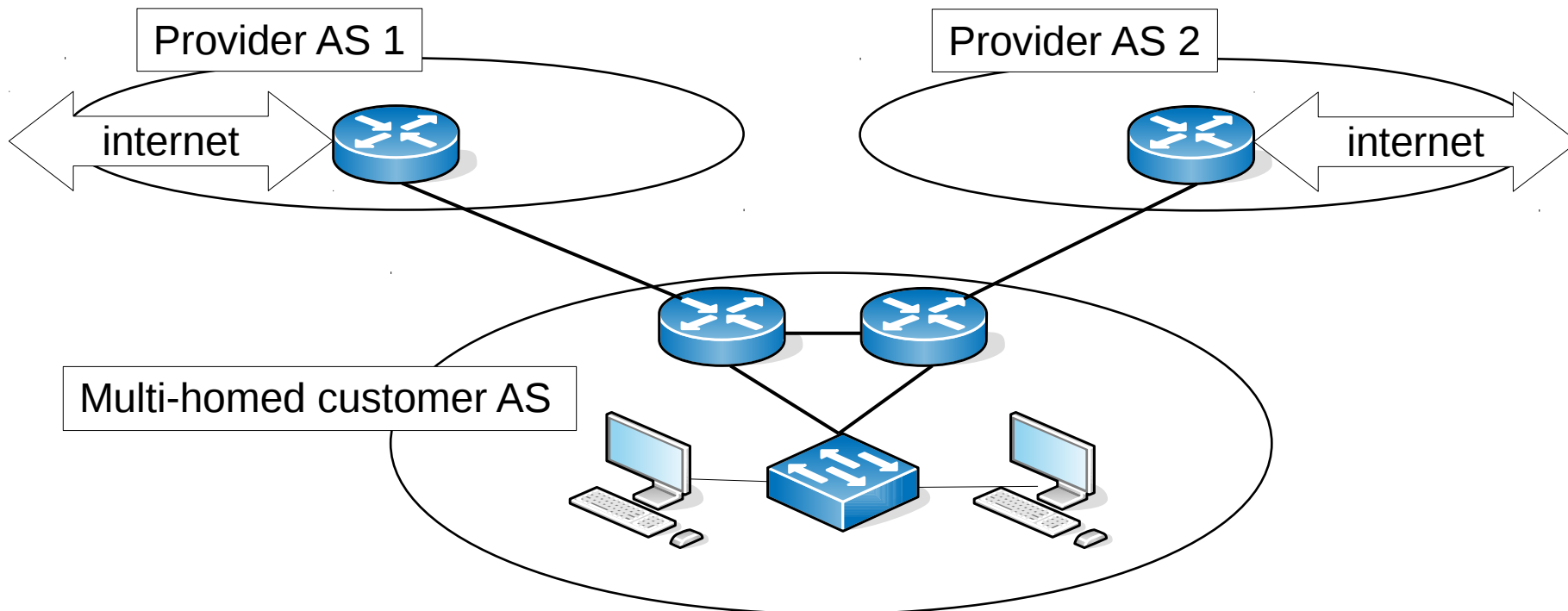
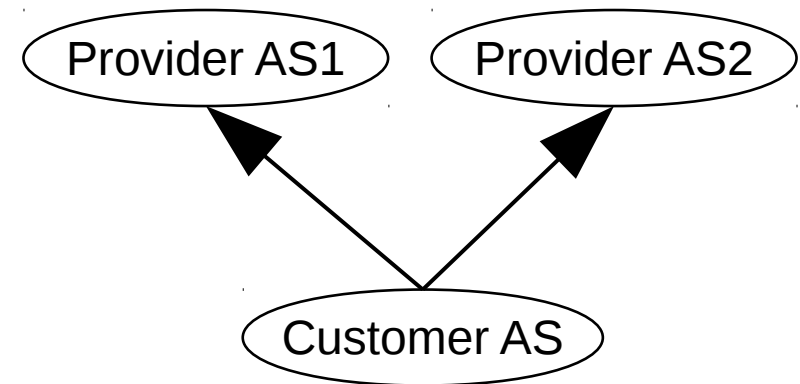
- A SH AS internetkapcsolata megszűnik, ha a szolgáltatója elérhetetlenné válik
- **Multi-homed (MH) AS:** olyan AS, amely több szolgáltatótón keresztül csatlakozik az internetre



MH AS: sematikus jelölés

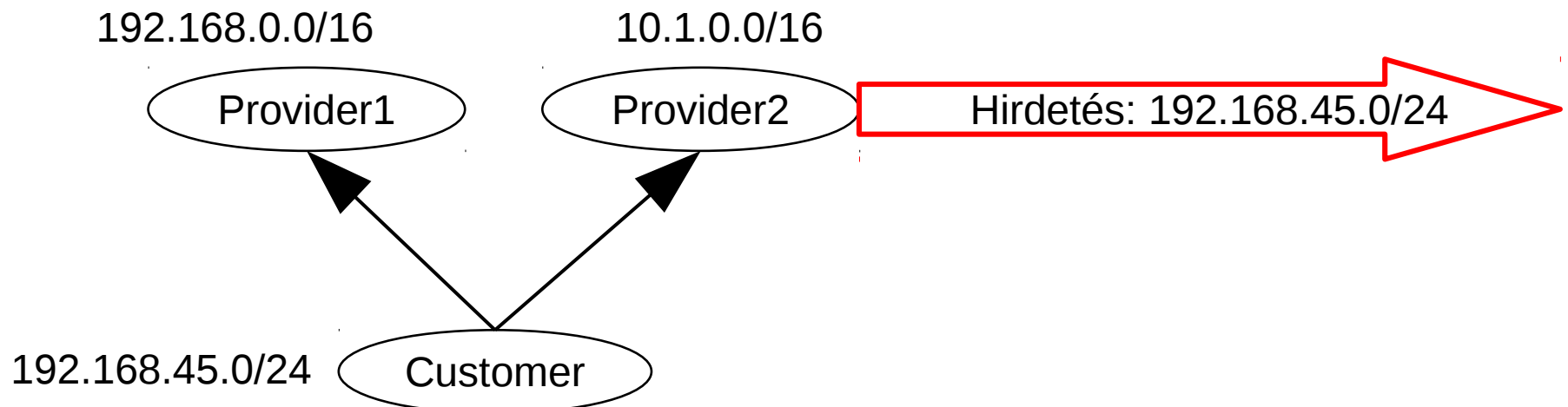
- Az előfizető AS-tól **nyíl** az összes szolgáltató AS felé
- A nyíl iránya megegyezik a cash-flow (\$\$) irányával

Sematikus jelölés



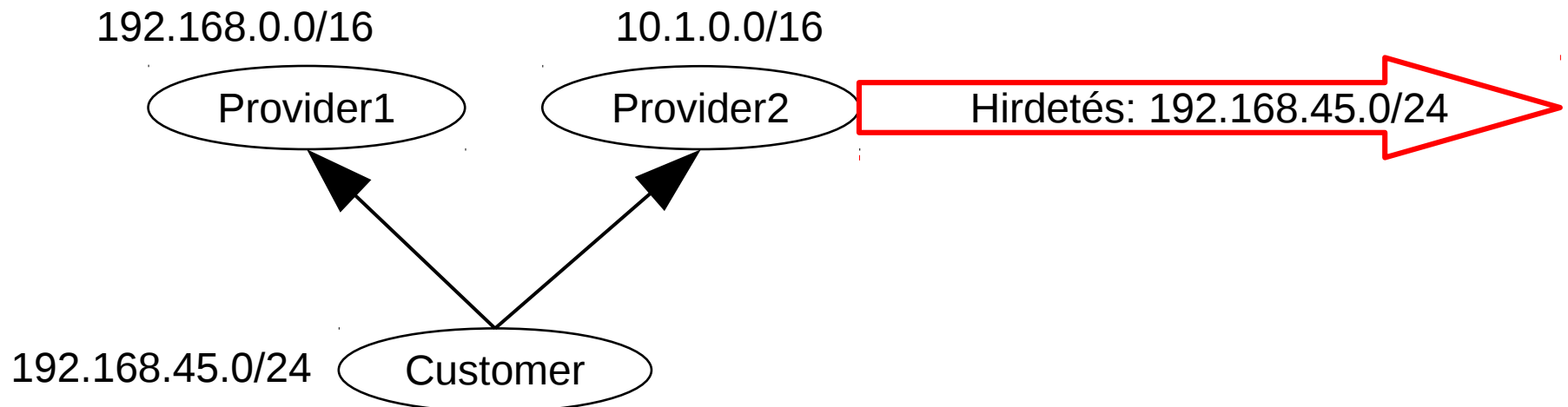
Multi-homed AS címzése: PA

- **Provider Aggregatable Addressing:** a MH AS az egyik szolgáltató alhálózati prefixét kapja
 - ettől még a másik szolgáltatónál nem aggregálható az előfizető címtartománya
 - külön hirdetni kell



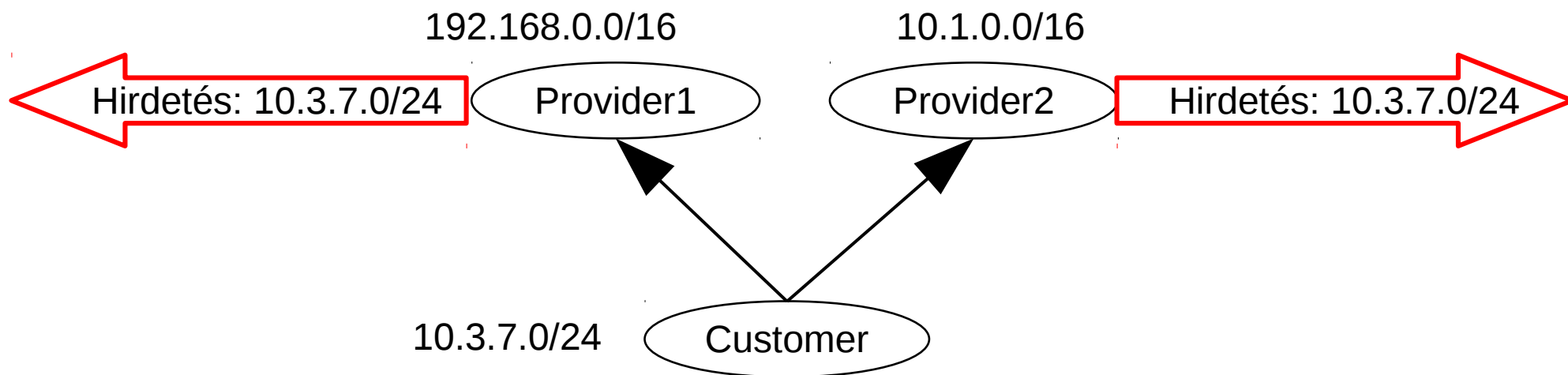
Multi-homed AS PA címzése: példa

- A példában a `192.168.45.0/24` prefix nem aggregálható a `Provider2` címtartományába
- Ezért `Provider2` kihirdeti, hogy rajta keresztül (is) elérhető a `192.168.45.0/24` prefix
- Plusz FIB bejegyzés az egész Interneten!



Multi-homed AS címzése: PI

- **Provider Independent Addressing:** az előfizető címtartománya nem része egyik szolgáltató címtartományának sem
 - mindkét szolgáltató külön hirdeti a PI prefixet
 - plusz egy FIB bejegyzés az interneten

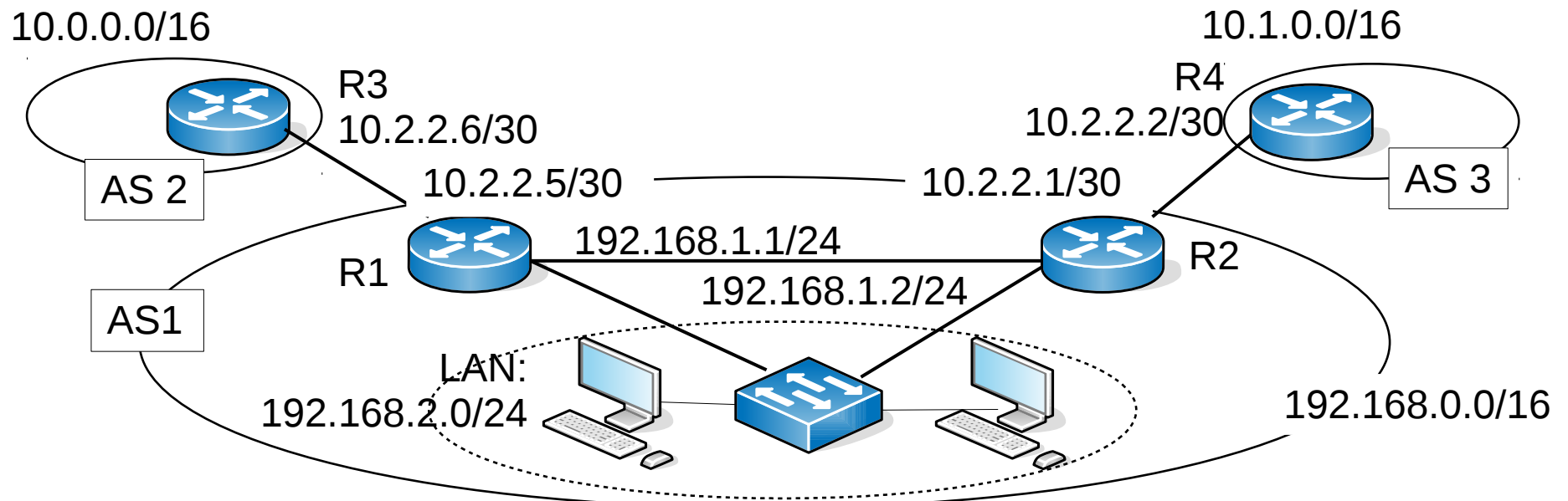


PA versus PI címzés

- PA címzés esetén az előfizető a szolgáltatóhoz van kötve: minden belső IP interfésze a szolgáltató címtartományából kap címet
 - drága szolgáltatót váltani (**renumbering**)
 - összes hosztot/routert újra kell konfigurálni
- PI címzésnél a szolgáltatóváltás egyszerűbb
 - sokszor SH AS-ek is igényelnek PI prefixet
 - szinte sosem kapnak, mert a plusz FIB bejegyzés az internet többi részén túl „drága”

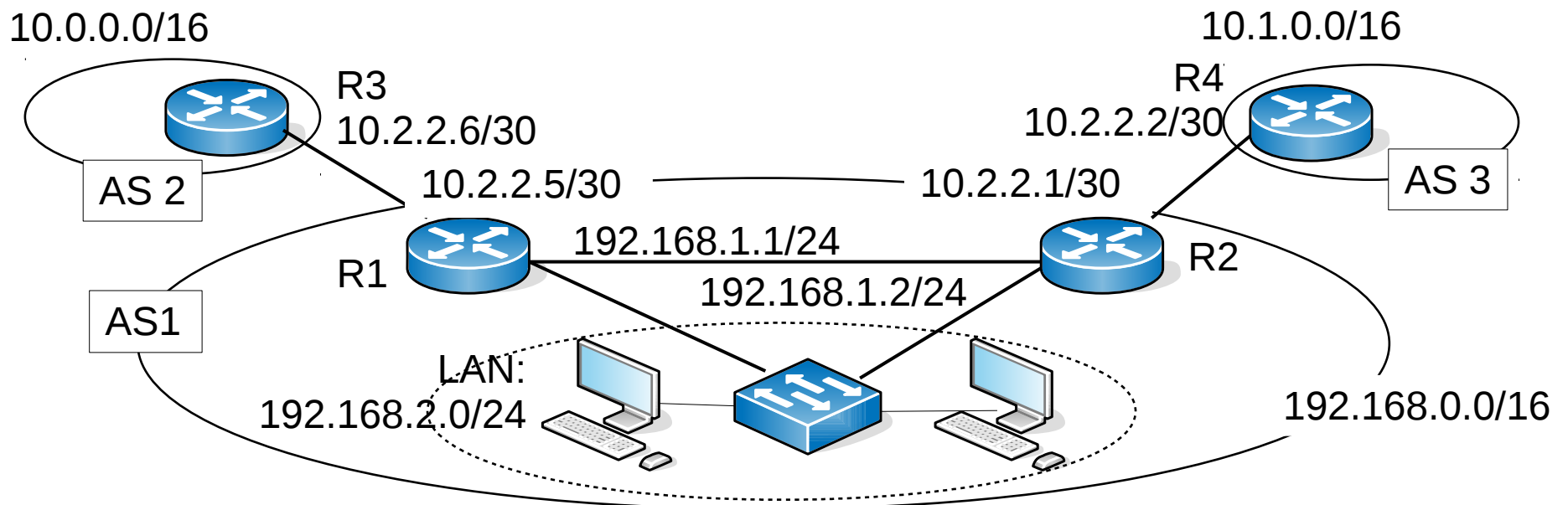
MH tranzit: példa

- **3 AS:** AS1 (címtartomány: 192.168.0.0/16), AS2 (10.0.0.0/16) és AS3 (10.1.0.0/16)
- AS1: két határrouter (R1 és R2) és egy LAN (címtartomány: 192.168.2.0/24)



MH tranzit: példa

R1 FIB		
Prefix	Next-hop	Megjegyzés
0.0.0.0/0	10.2.2.6	Default gateway R3 felé (AS2 az elsődleges szolgáltató)
10.0.0.0/16	10.2.2.6	AS2 forgalma R3 felé (opcionális default gateway miatt)
10.1.0.0/16	192.168.1.2	AS3 forgalma a másik határrouteren (R2) keresztül
192.168.2.0/24	local	a LAN lokális, a két hosztra nem kell külön bejegyzés

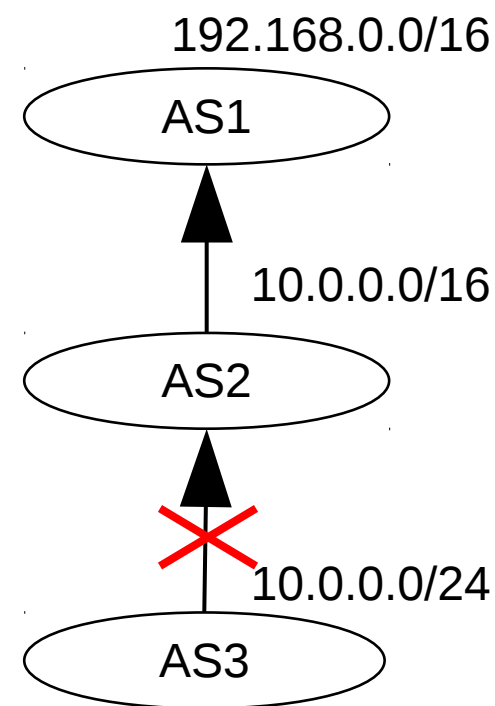


Multi-homing: értékelés

- Multi-homed kapcsolódás előnyei:
 - nagyobb rendelkezésreállítás
 - terhelésmegosztás a szolgáltatók közt
 - ároptimalizálás, jobb ártárgyalási pozíció
- Hátrányai:
 - bonyolultabb konfiguráció: dinamikus routing a határrouterekre (hiba esetén a másik szolgáltatóra kell kapcsolni)
 - multi-homed AS prefixe megjelenik minden Internet router FIB-jében: skálázhatatlan!

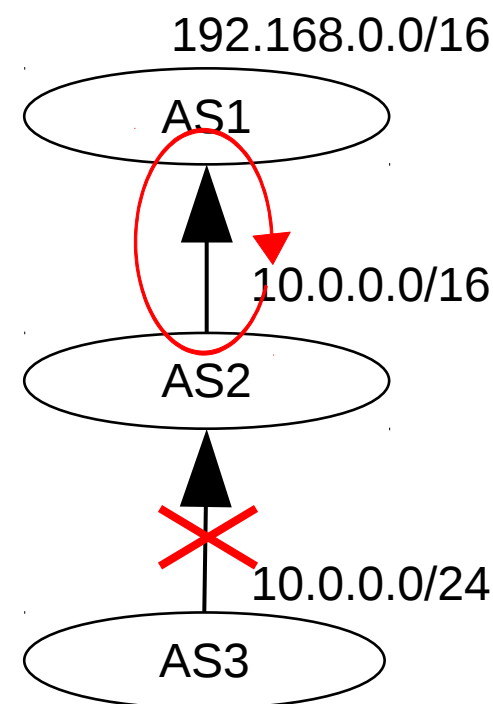
Tranzit: a default gateway veszélyei

- Több szolgáltatóból álló hierarchia esetén **forgalomtovábbítási hurok (loop)** jöhet létre
- Például ha AS1 szolgáltató AS2 számára PI prefixen, és AS2 szolgáltató AS3 számára PA prefixen
- Legyen AS2 default gateway-e AS1
- Tegyük fel, hogy AS1 egy csomagot kap AS3-beli címmel és az AS2-AS3 access link meghibásodik



Tranzit: a default gateway veszélyei

- AS1 továbbítja a csomagot AS2-nek
- AS2-nek nincs érvényes FIB bejegyzése AS3 prefixére, mert az AS2-AS3 link meghibásodott
- AS2 default gateway-re továbbít: a csomag visszakerül AS1-hez
- Amely ismét AS2-nek küldi azt
- Csomagtovábbítási hurok
 - csomag eldobódik, amint a TTL lejár

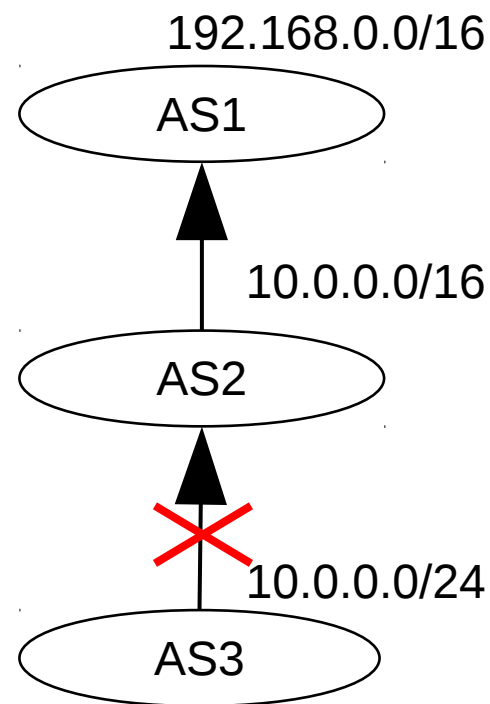


Tranzit routing: hurok elkerülése

- Sok módon elkerülhető, de legegyszerűbb, ha minden AS betartja az alábbi szabályt:

Ha egy AS tranzit szolgáltatást nyújt más AS számára, akkor nem tesz default gateway-t a FIB-be

- Nincs route: a csomag rögtön eldobódik, nem kerül hurokba
- Minden IPv4 prefix tárolva a FIB-ben
- **Default Free Zone:** azon routerek összessége, melyeknek nincs default gateway-ük



AS-ek mérete

- Egy AS jelentősége arányos azoknak az AS-eknek a számával, amelyek tőle vesznek tranzit szolgáltatást (**customer cone**)
- Az előfizetők száma helyett még jobb metrika az **AS által meghirdetett címtartomány nagysága** (hány IP címből áll a tartomány)
 - mivel az összes előfizetőjének a címtartományát meg kell hirdetni
 - a fenti szám jól jellemzi az AS-hez tartozó „customer cone” méretét
 - ezt nevezzük az **AS méretének**

A tranzit árazás

- A szolgáltatás minőségét az **SLA (Service Level Agreement)** rögzíti: formális szerződés
 - rendelkezésreállás (availability): például az idő 99.99%-ban elérhető a szolgáltatás
 - sebesség (rate), csomagvesztés (packet loss) és késleltetés (delay, delay variation)
 - monitoring: a felek hogyan mérik, hogy a szerződésnek megfelelő-e a szolgáltatás
- **Árazás:** sávosan a szolgáltató–előfizető AS linken áthaladó forgalom sávszélessége után

A tranzit árazás

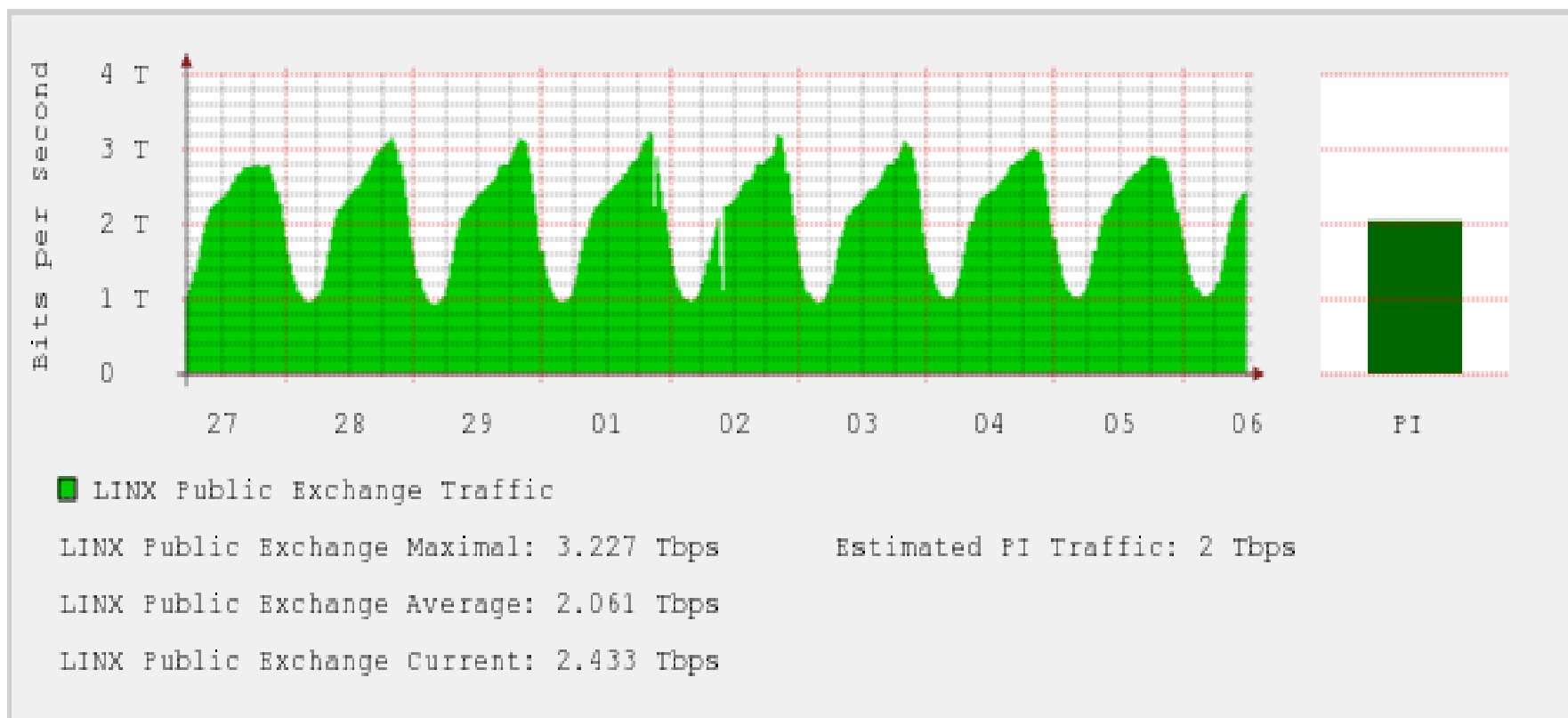
- **Commit:** előfizető várható havi adatsebessége

Commit rate	Egységár	Minimum fizetendő
10 Mbps	12\$ / Mbps	120 \$/hónap
100 Mbps	5\$ / Mbps	500 \$/hónap
1 Gbps	3.5\$ / Mbps	3 500 \$/hónap
10 Gbps	1.2\$ / Mbps	12 000 \$/hónap
100 Gbps	0.7\$ / Mbps	70 000 \$/hónap

- Pl. a „10 Mbps/12\$” commit választása esetén az előfizető akkor is kifizeti a 120\$-t, ha a sebessége kisebb volt, egyébként 12\$/Mbps
- De nagyobb commit-on kisebb az egységár!

A tranzit forgalom mérése

- A szolgáltatók forgalma órás, napi, heti, stb. skálákon tág határok közt változik



A tranzit forgalom mérése

- **95th percentile (95/5)** mérési módszer
 - 5 percenként sávszélességmérés a linken
 - hónap végén nagyság szerint növekvő sorba rendezik a mért sávszélességadatokat
 - mérési hibák eldobása (túlcsordulás, reboot)
 - 95th percentile: a legkisebb mért adat, amely már nagyobb vagy egyenlő, mint a minták 95%-a
- Másképp: az idő 95%-ában a forgalom a kapott érték alatt volt (a ritka burst-ök levágása)

A tranzit árazás: példa

- Például a $[1, \dots, 1, 100]$ 100 mérésből álló sorozathoz tartozó 95th percentile: 1 Mbps
- De a $[1, \dots, 1, 15, 16, 16, 17, 21, 40]$ 100 elemű sorozathoz már 15 Mbps az érték (a 95. elem a nagyság szerint rendezett listában)
- Ez utóbbihoz a 10Mbps commiton a havidíj

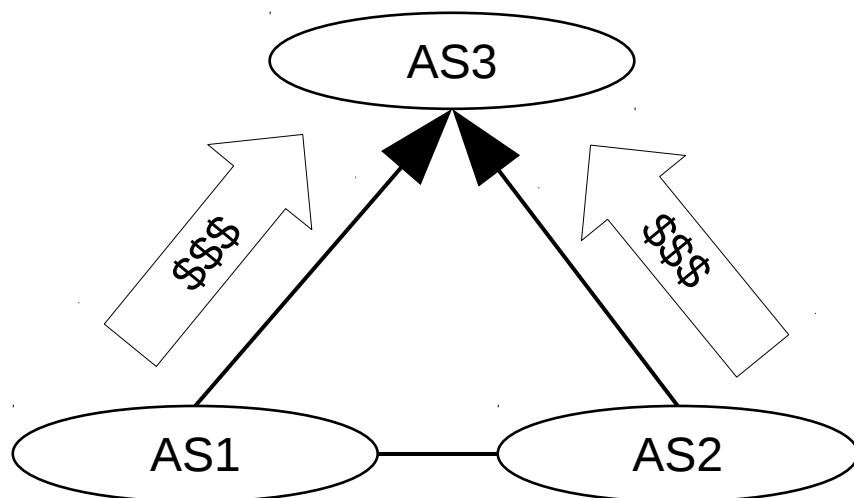
$$12 \text{ \$/Mbps} * 15 \text{ Mbps} = 180\text{\$}$$

- De 100 Mbps commiton 500\$, mert a minimum összeget a commit után fizetni kell!

$$\text{commit rate (100\$)} * \text{egységár (5\$/Mbps)} = 500\text{\$}$$

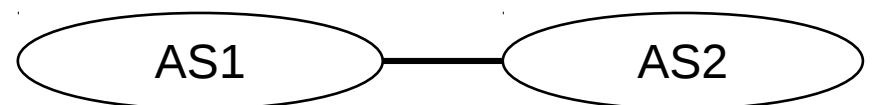
A peer AS-AS kapcsolat

- Tegyük fel, hogy az alábbi konfigurációban AS1 és AS2 mindketten előfizetői AS3-nak, és nagyjából azonos méretűek
- Ekkor olcsóbb, ha egymás forgalmát ingyen kicserélik egymás közt, AS3 közbeiktatása nélkül: **peer** kapcsolat



- A peer kapcsolat jelén nincs nyílhegy, jelezve, hogy nincs cash-flow

Sematikus jelölés



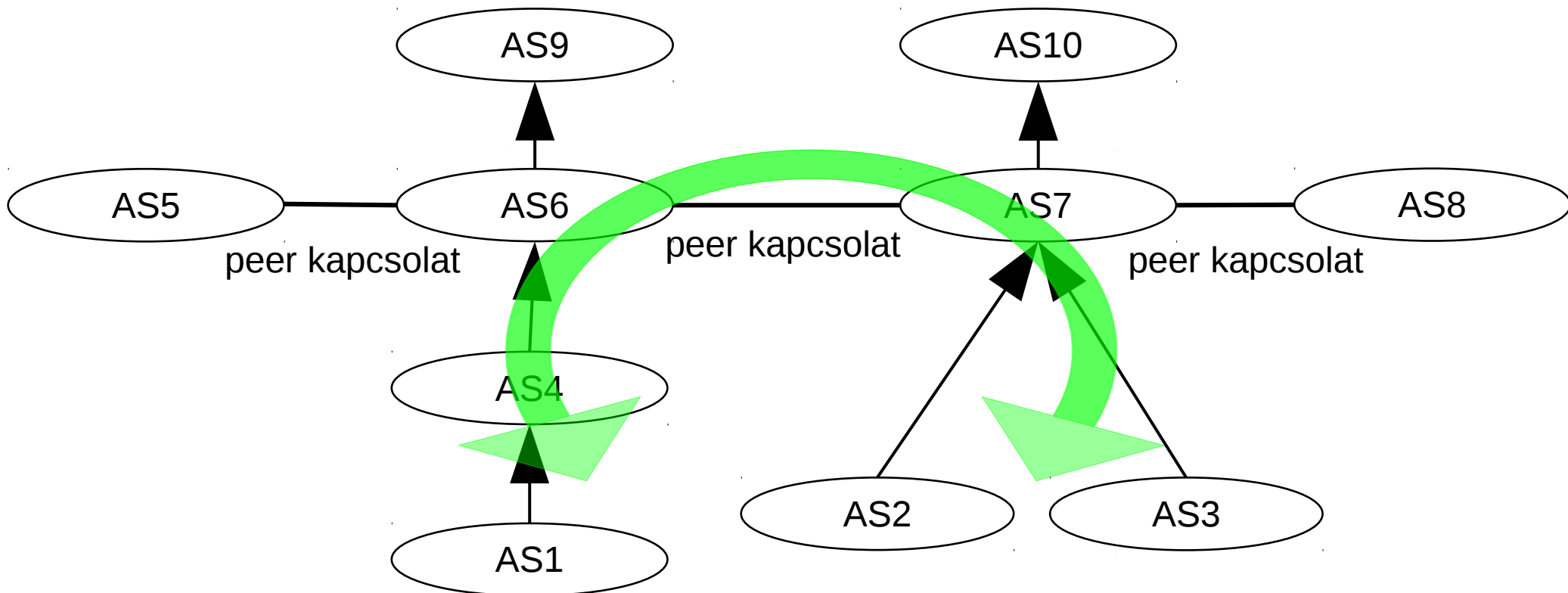
A peer AS-AS kapcsolatok

Internet Peering is the business relationship whereby companies reciprocally provide access to each others' customers.

(DrPeering)

A peer AS-AS kapcsolat

- **Peer kapcsolat:** bilaterális egyezés két AS között, hogy egymás, illetve egymás tranzit előfizetői között térítés nélkül továbbítják a forgalmat



Peering megfontolások

- A peer kapcsolat kvázi ingyen van (persze a link fenntartása sok pénzbe kerülhet)
- Még sincs minden AS között peer kapcsolat
- Nagyobb AS például nem peer-el egy kisebbel
 - hiszen így ingyen hozzáférést enged a saját előfizetőihez
 - ehelyett tranzitot is szolgáltatna, pénzért
 - profitveszteség megengedhetetlen egy ilyen kompetitív piacon

Peering policy

- Minden AS megszabja, milyen feltételek mellett hajlandó peer kapcsolatot létrehozni
- **Selective peering:** szigorú peering policy
 - nagyjából hasonló méretű customer cone
 - szimmetrikus forgalmi igény a két AS közt
 - több, földrajzilag eltérő POP-on is jelen van
 - 24x7 support
- **Open peering:** bárkivel peer kapcsolat
- Lásd: PeeringDB

Peering wars

- **Döntési helyzet: hozzunk-e létre peer kapcsolatot egy AS-sel?**

Peer kapcsolat	Nem peerel
csökkentheti a tranzit díjakat (a peer linken haladó forgalom után nem kell fizetni a tranzitnak)	profitveszteséget okozhat (ahhoz képest, mintha ha a másik AS inkább előfizetőnk lenne)
csökkentheti a késleltetést (a peer linken közvetlen forgalomcsere, a tranzit közbeiktatása nélkül)	kompetitív hátrány: a másik AS hozzáfér az előfizetőinkhez

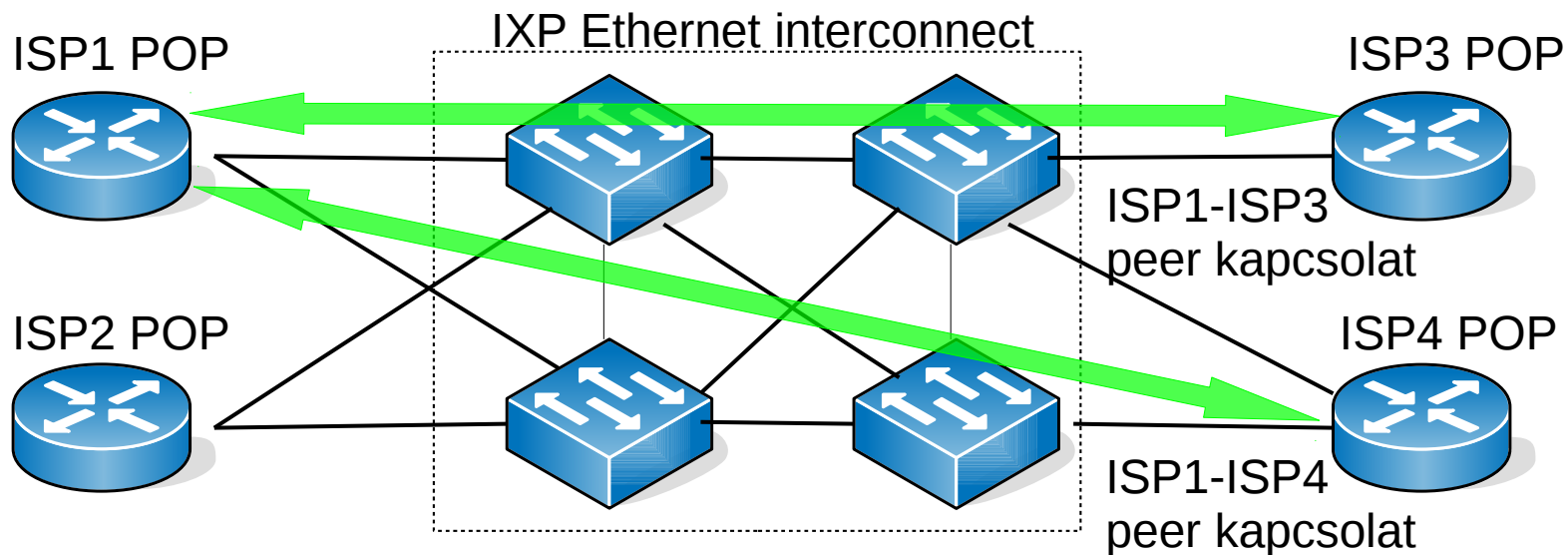
- **Messze a legtöbb vitát kiváltó konfliktus az interneten**

IXP

- A peer kapcsolatok nagy része IXP-kben létesül
- **Internet eXchange Point (IX/IXP):** speciális infrastruktúra, amelynek célja hogy AS-ek tömegesen peer kapcsolatot létesítsenek
- Tipikusan egy épület egy hűtött terme, amelyben az ISP-k POP-okat hozhatnak létre
- Az IXP biztosítja, hogy bármelyik POP bármelyik másikkal kapcsolódni tudjon
- Ha egy ISP jelen van egy IXP-n, akkor onnan nagyon olcsó új peer kapcsolatokat kiépíteni

IXP

- Az IXP lehet for-profit (USA) vagy non-profit (Európa) vállalkozás
- Célja minél egyszerűbb kapcsolódási pontot nyújtani, minél több ISP-t vonzani
- A kicserélt forgalom után díjat szed be



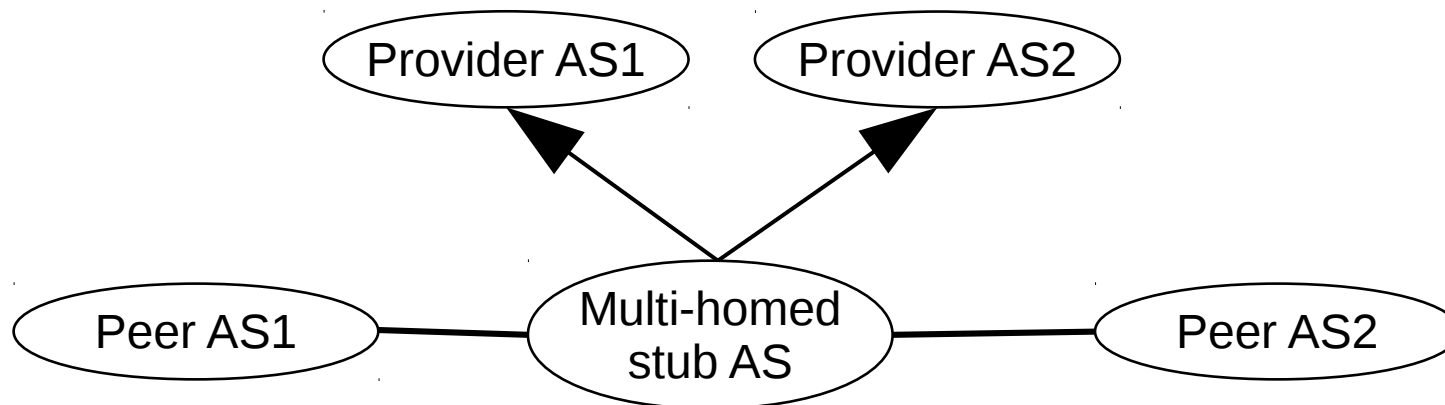
IXP

- Több száz AS-t szolgálhat ki, az átvitt forgalom akkora lehet, mint a legnagyobb ISP-ké
 - DE-CIX (Frankfurt, Hamburg, Munich): 600+ ISP, 1.7 Tbps átlagos forgalom
 - AMS-IX (Amsterdam Internet Exchange), LINX (London Internet Exchange), Equinix
 - BIX: Budapest Internet eXchange



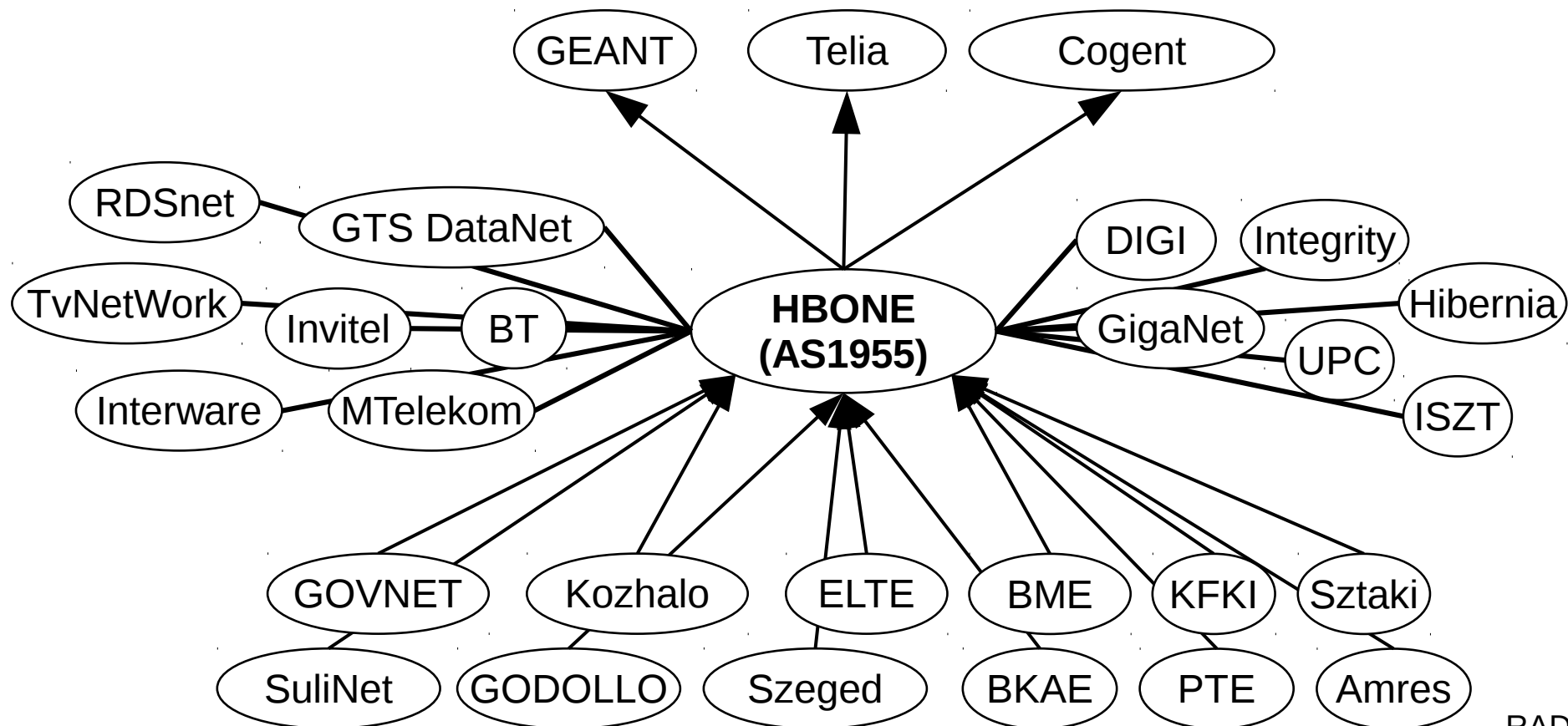
Stub versus tranzit AS-ek

- **Stub AS:** olyan AS, amely nem végez tranzit szolgáltatást más AS számára
- Az interneten jelenleg aktív kb. 40 ezer AS-nek becslések szerint nagyjából fele/harmada stub
- **Tranzit AS:** értelemszerűen, „nem stub” AS



HBONE (AS1955)

- Magyar akadémiai számítógéphálózat: felsőoktatás, kutatás-fejlesztés, könyvtárak, közgyűjtemények, kormányzati szervek, közintézmények



Az internet AS-szintű struktúrája

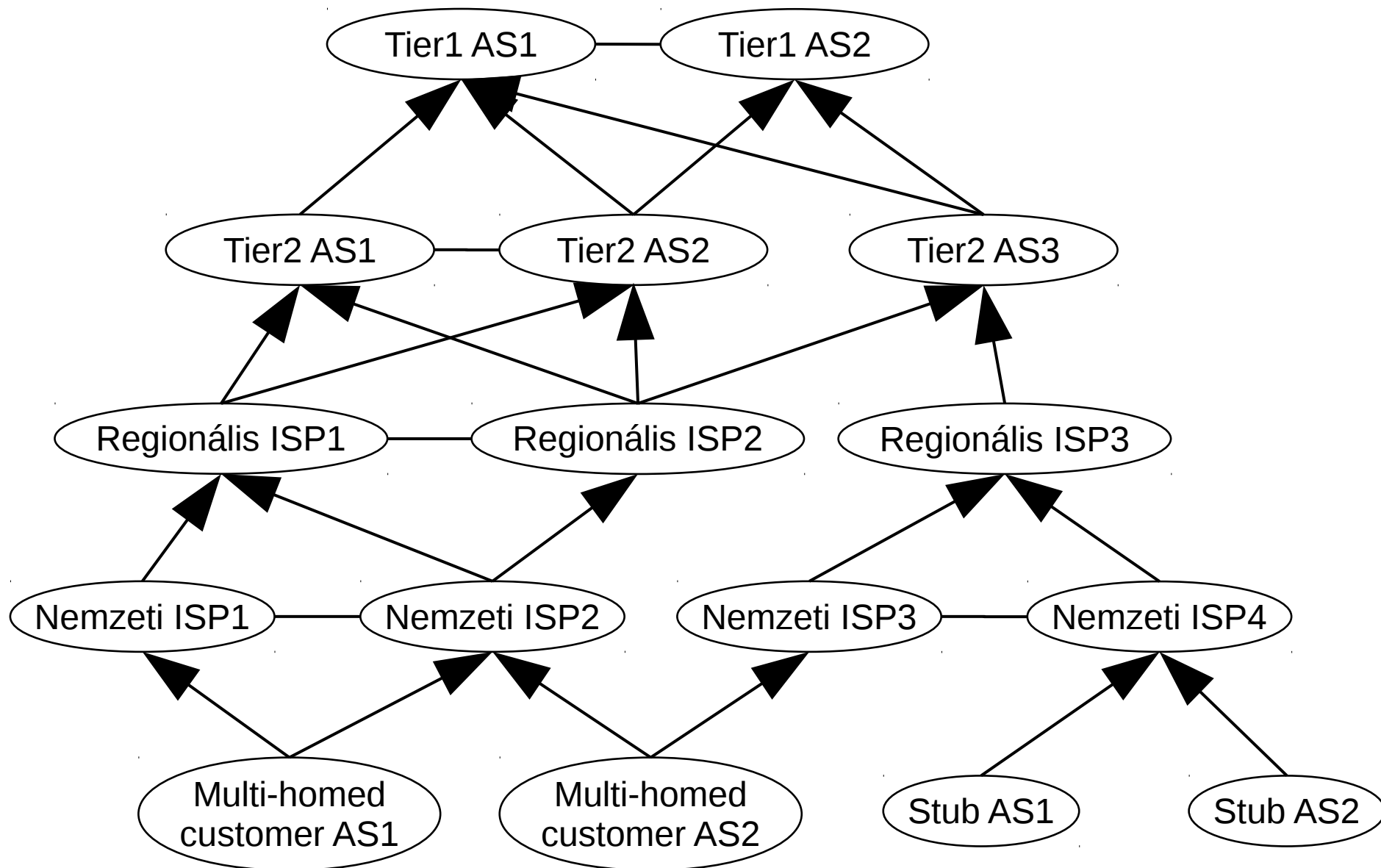
Tranzit hierarchia: a Tier-ek

- Egy előfizetői AS tranzit szolgáltatást nyújthat további AS-eknek, és így tovább
- **Tier 1:** globális ISP, amelynek nincs már további szolgáltatója (nincs **upstream AS**)
 - globális: a világ bármely részén levő két előfizetőt össze tud kapcsolni
 - Tier1 AS-ek egymás közt **full peer-mesh**
 - 12 Tier1 AS: AT&T, CenturyLink, Cogent, GTT, Deutsche Telekom, Level3, NTT, Sprint, Tata, Seabone, TeliaSonera, Verizon, XO

Tranzit hierarchia: a Tier-ek

- **Tier 2:** Tier 1 AS előfizetője (gyakran szintén globális hálózattal rendelkezik)
- **Regionális szolgáltató:** adott földrajzi régión belül szolgáltató (pl. Comcast: USA, Orange: Európa)
- **Nemzeti szolgáltató:** egy ország piacán nyújt internet szolgáltatást
- A hierarchia alján: **single-** és **multi-homed előfizetők, stub AS-ek**

Tranzit hierarchia: a Tier-ek



Internet: ismeretlen terep

- A tranzit-peer modell az AS-AS kapcsolatok csak kb. 70%-át fedi le
 - **paid peering**: peering IXP-n, de tranzitdíjért
 - **sibling**: kölcsönös tranzit szolgáltatás
 - és létezik sok, általunk ismeretlen policy is
- Az AS-AS üzleti kapcsolatok titkosak!
 - ismerni egy AS üzleti stratégiáját nagy előny
 - néhány AS nyilvánosságra hozza: IRR
 - vagy következtetés traceroute adatokból

Internet: változó terep

- A Tier-felosztás csak hozzávetőleges
 - sok Tier1 közvetlenül szolgáltat nemzeti ISP-knek, sőt, stub AS-eknek is
 - nem csak azonos tier-en levő AS-ek közt lehet peer kapcsolat!
 - néha a Tier1-ek is megszakítják a peer kapcsolatot (**depeering**) → peering wars
- **Flattening:** az internet a kezdeti szabályos tranzit AS-hierarchiából egyre inkább a peer full-mesh irányában fejlődik
- De a pontos AS-szintű topológia nem ismert!!!