

# IP anycast

Jákó András

[jako.andras@eik.bme.hu](mailto:jako.andras@eik.bme.hu)

BME TIO

# Tematika

Mi az IP anycast?

Hogy működik?

Mire használható?

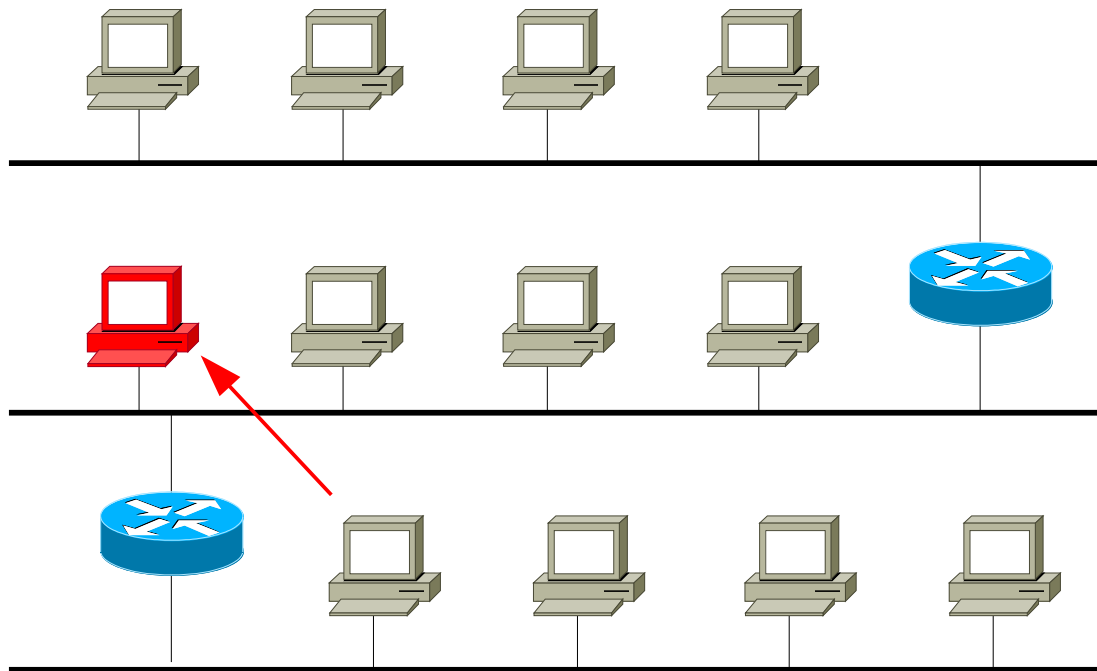
Alkalmazási példa

# IP ...cast

- IP csomagtovábbítási módok a célállomások szerint
  - unicast
  - multicast
  - broadcast
  - anycast

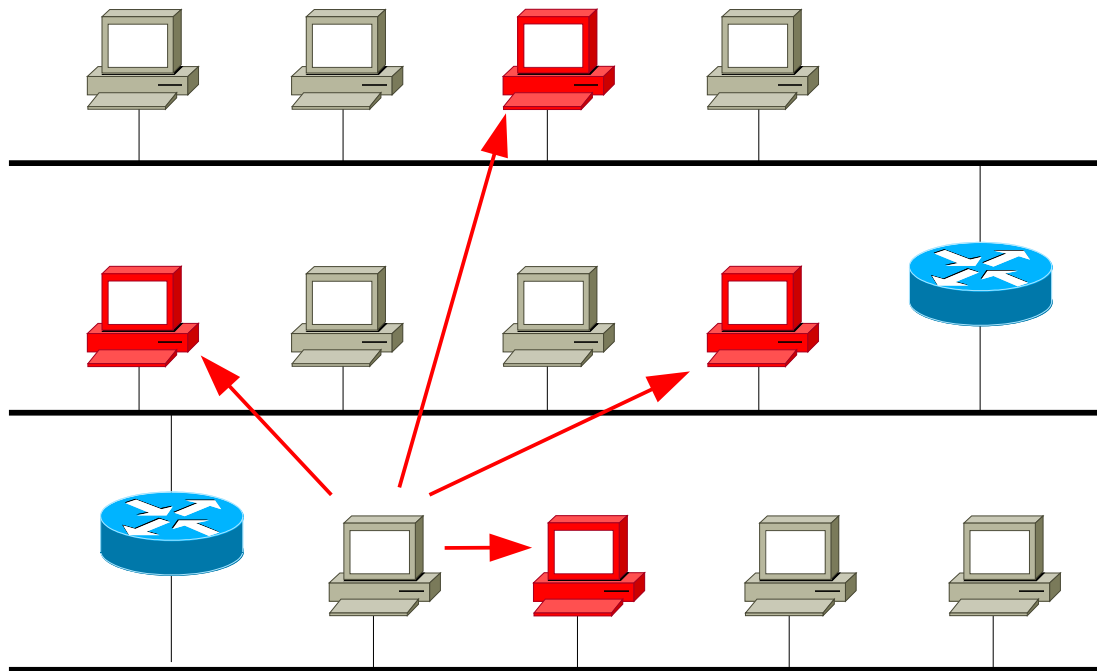
# IP unicast

- egy célállomás



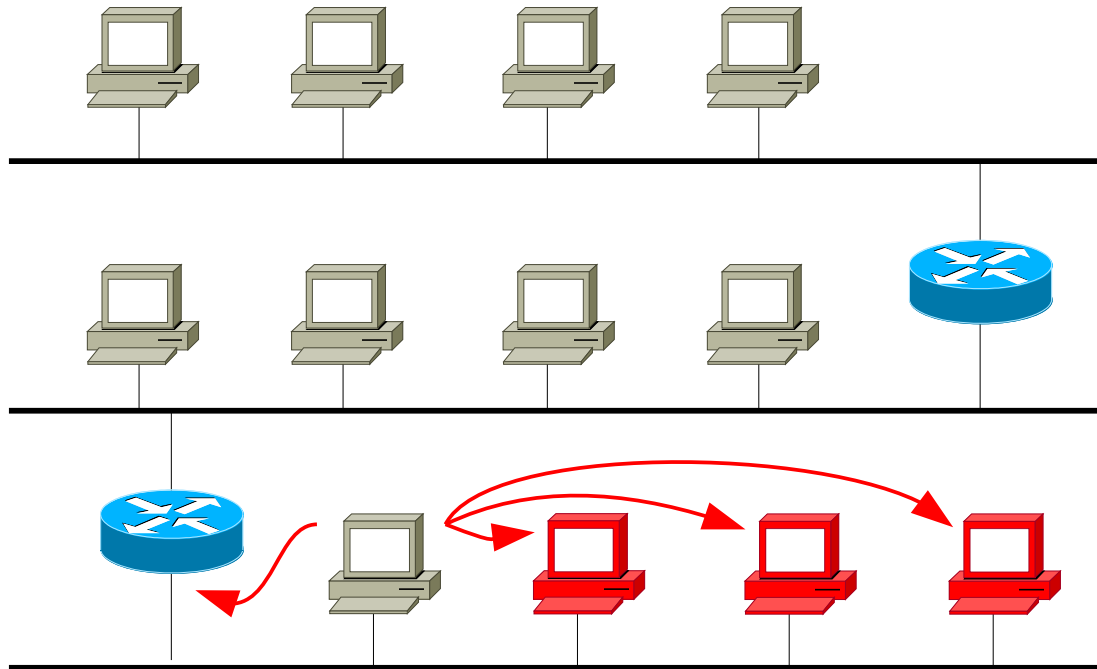
# IP multicast

- több célállomás



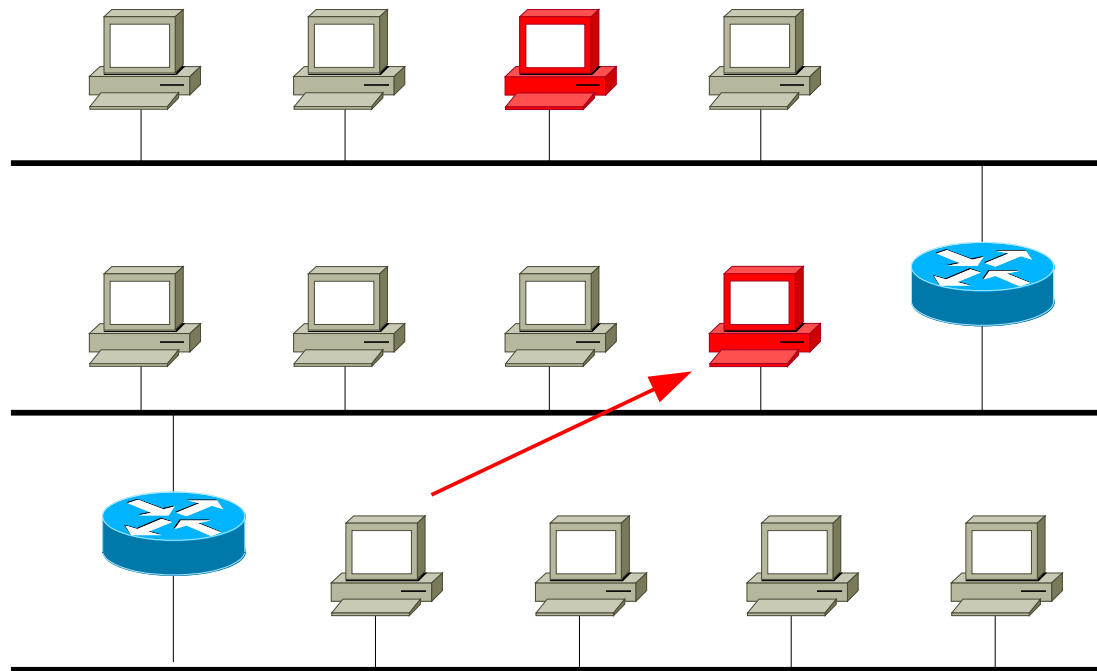
# IP broadcast

- minden állomás
  - a saját subneten/linken



# IP anycast

- egy célállomás
- több lehetséges közül bármelyik



# Tematika

Mi az IP anycast?

Hogy működik?

Mire használható?

Alkalmazási példa



# Alapvető jellemzők

- több kiszolgáló példány
  - (általában) azonos szolgáltatást nyújtanak
- a példányok IP címe azonos
- a kliens változatlan módon működik
  - nem is tud róla, hogy több példány van
- a kiszolgáló példányok alapműködése változatlan
  - IP csomagok vétele és küldése változatlan
  - de kellhet egy sor kiegészítés, a megvalósított szolgáltatástól függően
- az adott klienst kiszolgáló példány időben változhat
- a példányok helye szerint kétféle: subnet/global

# IPv6 „subnet” anycast

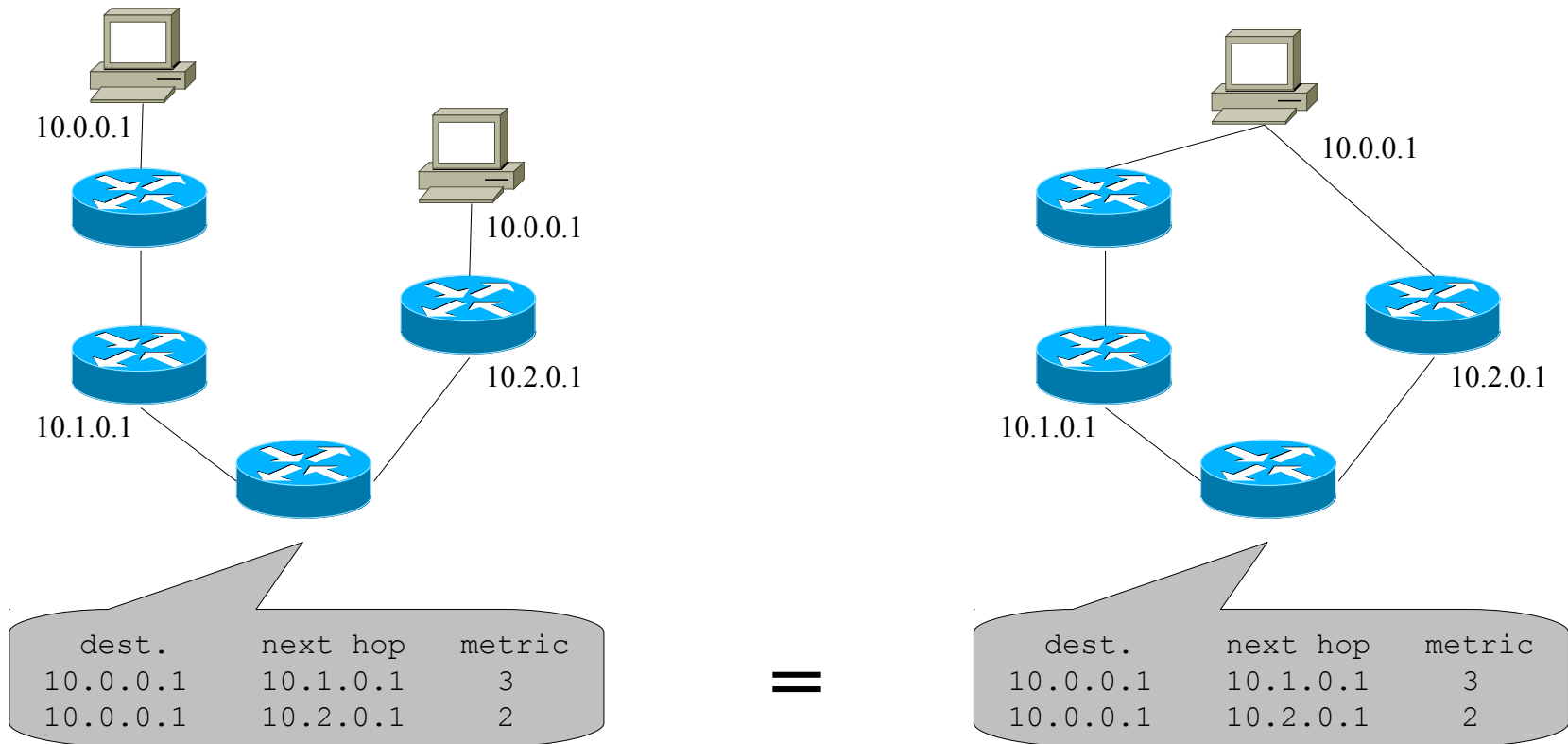
- csak IPv6
- egy linken vannak az azonos IP című példányok
- a Neighbor Discovery választ a példányok közül
  - a válasz Neighbor Advertisementet minden példány véletlenszerű késleltetéssel küldi
  - a válaszban az Override bit nincs beállítva
    - a korábban érkezett válaszokat nem írja felül
  - az első válaszoló MAC címe kerül a Neighbor Cache-be
  - a szokásos módon használt Neighbor Unreachability Detection segítségével törlődik a kiesett példány a kliens Neighbor Cache-éből

# Globális IP anycast

- a példányokat (egy/több) router választja el egymástól
  - a példányok az Interneten bárhol lehetnek
  - az anycast cím (prefix) megjelenik a routing táblában, példányonként
- a rendes unicast routing választ a példányok közül
  - a routerek változatlan módon működnek
- a klienst a hozzá legközelebbi példány szolgálja ki
  - legközelebbi: a routingban használt mérték szerint

# Anycast routing

- a routerek úgy látják, mintha
  - egy példány lenne
  - ami több útvonalon is elérhető

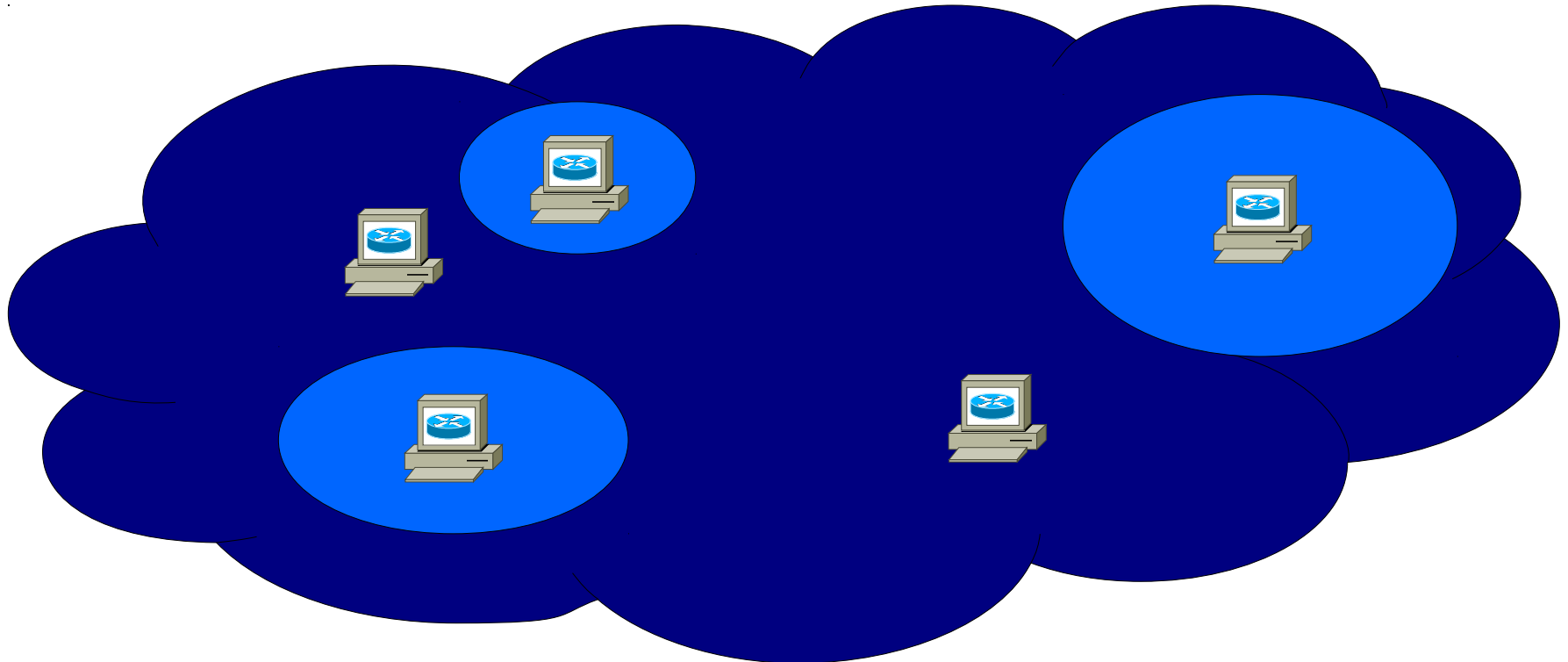


# Anycast példányok megvalósítása

- egy loopback interface kapja meg az anycast címet
- a példánynak megvan az egyedi unicast címe is
  - felügyelethez, karbantartáshoz
  - példányok megkülönböztetését igénylő műveletekhez
  - állapotfenntartást igénylő műveletekhez
- a host valamilyen routing protokollt beszél
  - a szolgáltatás működőképessége esetén jelenti csak be az anycast prefixet
    - lehetőleg automatizáltan
  - ezáltal könnyen ki is vehető az aktív példányok közül pl. karbantartásra

# Példányok által ellátott „terület”

- a prefix hirdetésével szabályozható
  - a unicast routingban szokásos módszerekkel
- szokásos a lokális és globális példányok keverése
  - **lokális**: az anycast prefix hirdetése korlátozott
  - **globális**: az anycast prefix hirdetése az Interneten bármeddig eljuthat



# Tematika

Mi az IP anycast?

Hogy működik?

Mire használható?

Alkalmazási példa

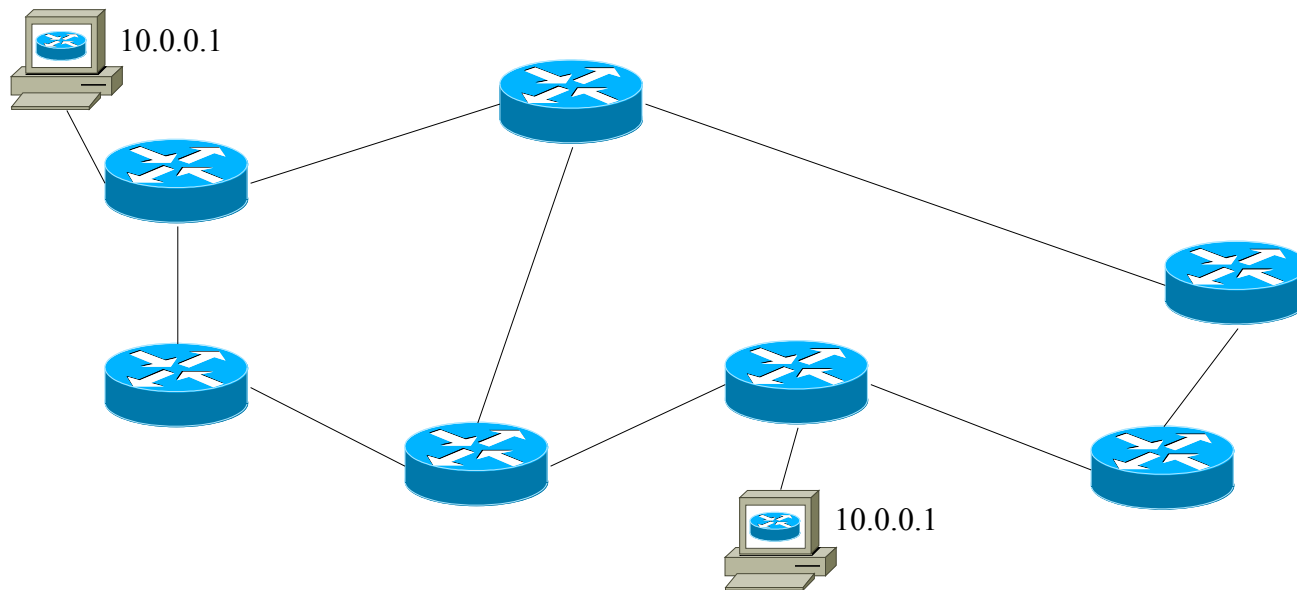
# Korlátok

- idővel változhat, hogy a kliens melyik példánnyal kommunikál
  - rövid üzenetváltásokra megbízható
    - pl. DNS lekérdezés
  - állapotmentes kommunikációra hosszan is jó
    - pl. 6to4 relay
- lehetőségek hosszabb kommunikáció esetén
  - anycast csak a felderítési szakaszban
  - átirányítás a példány unicast címére (pl. webes alkalmazásnál)
    - állapotárolás csak a kliensnél



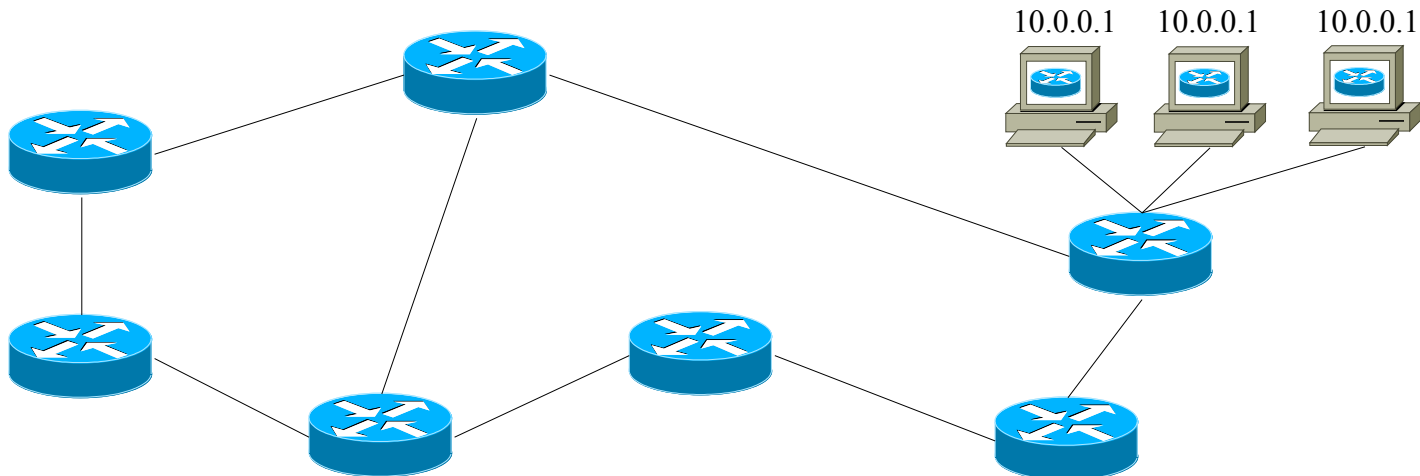
# Terheléselosztás

- távoli példányok
  - a kliensek hálózati helyük szerint oszlanak el a példányok közt
    - annyira stabil, amennyire a BGP topológia stabil



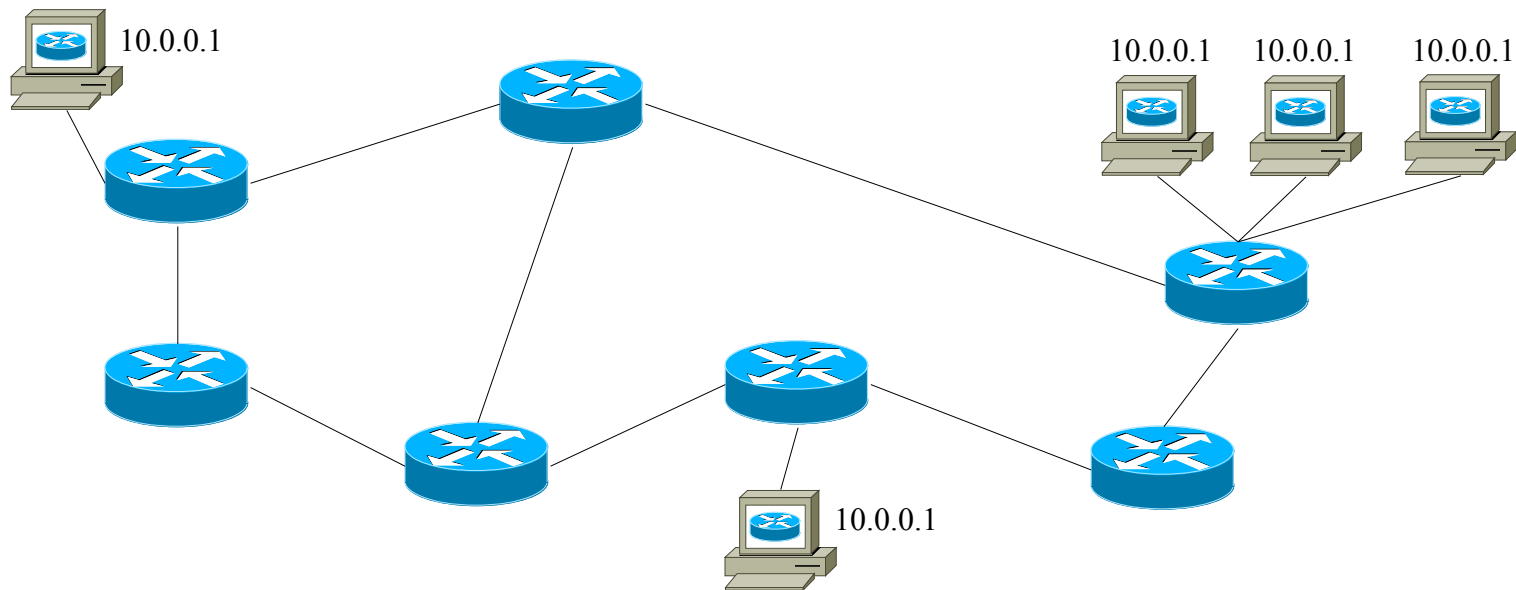
# Terheléselosztás

- távoli példányok
  - a kliensek hálózati helyük szerint oszlanak el a példányok közt
    - annyira stabil, amennyire a BGP topológia stabil
- cluster: egymáshoz közeli példányok, egy router mögött
  - a kliensek az elosztó router ECMP algoritmusára szerint oszlanak el a példányok közt
    - PPLB csak 1 kérdés-válasz esetén használható



# Terheléselosztás

- távoli példányok
  - a kliensek hálózati helyük szerint oszlanak el a példányok közt
    - annyira stabil, amennyire a BGP topológia stabil
- cluster: egymáshoz közeli példányok, egy router mögött
  - a kliensek az elosztó router ECMP algoritmusára szerint oszlanak el a példányok közt
    - PPLB csak 1 kérdés-válasz esetén használható
- ezek keveréke



# Védelem

- hibatűrés
  - egy kieső példány kliensei a második legközelebbi példányhoz fordulnak
    - ehhez el kell tűnnie a kiesett példány hirdetésének
- DoS támadás enyhítésére
  - a támadási terhelés megoszlik a példányok közt, nem bénítja meg mindet
    - ehhez meg kell maradnia a példányok hirdetésének

# Tematika

Mi az IP anycast?

Hogy működik?

Mire használható?

Alkalmazási példa

# DNS

- csak lekérdezésekhez
  - általában UDP, 1 kérdés 1 válasz
  - ritkán TCP, de az is rövid
- root nameserverek
  - 13-ból 10 anycast
  - root serverenként 2-70 példány (összesen ~240)
  - ellátott terület szerint globális és lokális példányok általában vegyesen
  - <http://www.root-servers.org/>
- AS112
  - RFC1918 címek reverse zónái
- stb.

# DNS: k.root-servers.net

**Global Nodes:** Amsterdam, NL • London, GB • Tokyo, JP • Miami, Florida, US • Frankfurt, DE



**Local Nodes:** Budapest, HU • Milan, IT • Helsinki, FI • Reykjavik, IS • Poznan, PL • Geneva, CH • Delhi, IN • Athens, GR • Doha, QA • Novosibirsk, RU • Abu Dhabi, AE • Brisbane, AU • Dar es Salaam, TZ

# DNS: k.root-servers.net

- AS2547, Müegyetem:

Tracing the route to k.root-servers.net (193.0.14.129)

```
1 tge2-2.sup720.bme.hbone.hu (152.66.0.126) 0 msec 0 msec 4 msec
2 193.6.27.61 [AS 1955] 0 msec 0 msec 4 msec
3 fe00.router.bix.k.ripe.net (193.188.137.130) [AS 5507] 0 msec 0 msec 0 msec
4 k.root-servers.net (193.0.14.129) [AS 25152] 0 msec 0 msec 0 msec
```

- AS5400, Barcelona:

Tracing the route to k.root-servers.net (193.0.14.129)

```
1 t2a1-p1-1.es-bar.eu.bt.net (166.49.163.38) 4 msec 0 msec 0 msec
2 t2c1-ge7-0.es-bar.eu.bt.net (166.49.143.43) 4 msec 0 msec 0 msec
3 t2c1-p2-0.nl-ams2.eu.bt.net (166.49.164.33) 28 msec 24 msec 28 msec
4 t2a4-ge6-0-0.nl-ams2.eu.bt.net (166.49.200.20) 24 msec 28 msec 28 msec
5 g00.router.ams-ix.k.ripe.net (195.69.144.240) [AS 1200] 24 msec 28 msec 28 msec
6 k.root-servers.net (193.0.14.129) [AS 25152] 28 msec 28 msec 24 msec
```



# Köszönöm a figyelmet!



az előadás diái:  
<http://splash.eik.bme.hu/>