Távközlő hálózatok és szolgáltatások

Forgalmi követelmények, hálózatméretezés

Csopaki Gyula
Németh Krisztián
BME TMIT

2013. nov. 19.

A tárgy felépítése

- 1. Bevezetés
- 2. IP hálózatok elérése távközlő és kábel-TV hálózatokon
- 3. VoIP
- 4. Kapcsolástechnika
- 5. Mobiltelefon-hálózatok
- 6. Forgalmi követelmények
- 7. Jelzésátvitel
- 8. Gerinchálózati technikák (Cinkler Tibor)

Forgalmi követelmények, hálózatméretezés

- Távbeszélő-hálózatok forgalmi jellemzése
- Csomagkapcsolt hálózatok forgalmi jellemzése
- Egy gondolat a teljesítőképességről

Távbeszélő-hálózatok forgalmi jellemzése

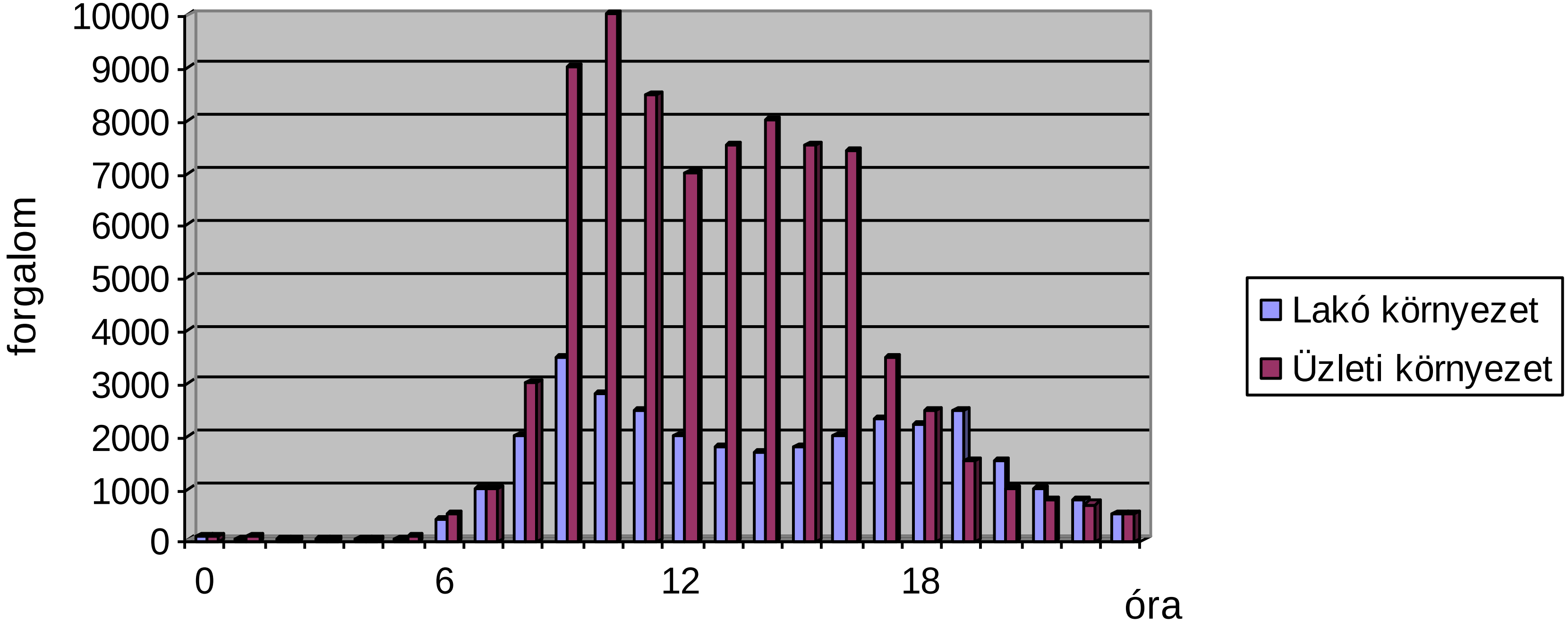
- Cél: hálózat méretezése
- Pl. 10 000 előfizető $\neq$ 10 000 áramkörre méretezett kapcsolóközpont és átviteli utak
- Cél pontosabban: *legkevesebb hány ák. kell, hogy a blokkolás adott érték alatt maradjon?*
  - Blokkolás: {túlterhelés miatt sikertelen hívások száma} / {összes hívás száma}
- Ehhez kell: forgalmi statisztikák
  - pl. az előfizetők mikor, milyen gyakran, milyen hosszan beszélnek

Távbeszélő-hálózatok forgalmi jellemzése

- Ehhez két leíró:
  - $X(t)$ – $[0,t]$ intervallumban beérkezett hívások száma
  - $Y(t, \tau)$ – a hívások tartásideje:
    - két dimenziós eloszlás sűrűségfüggvénye, ahol egy hívás t időben kezdődik és $t+\tau$ -kor végződik
  - $Y(t, \tau)$-ről feltételezzük, hogy:
    - független az előző kimenetelektől (OK),
    - és a felhasználótól (!),
    - de az időtől (t) függhet
- Adott a maximálisan elfogadható blokkolás
- Hány áramkör kell?
- Ez így túl nehéz, de nekünk elég egy egyszerűbbet megválaszolni

Távbeszélő-hálózatok forgalmi jellemzése

- Ugyanis a tapasztalat a *forgalomról* (=egyszerre hány hívás van egy központban) ilyen jellegű:



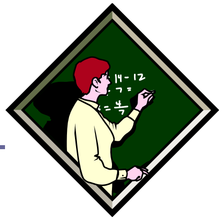
- „Worst case” méretezés: a „csúcsforgalomra” tervezünk
- Ekkor jó közelítéssel *X(t)* növekménye stacionárius: a forgalom kb. állandó
- Erre az időszakra mondhatjuk, hogy a tartásidő eloszlása is független az időtől (azonos)

Távbeszélő-hálózatok forgalmi jellemzése

- Ekkor jelentősen egyszerűsödik a modell.
- A megfigyelések alapján jó közelítés:
  - $X(t)$ -- Poisson folyamat. Várható érték=param.= λ
    - λ -- hívásgyakoriság [1/óra]
  - $Y(t, \tau) = Y(\tau) \leftarrow$ exponenciális eloszlás. Várható érték=1/param.= h
    - h -- átlagos tartásidő [perc] (!)
- A – forgalomintenzitás
  - $A = \lambda \cdot h$
  - A [1], de szokás Erl-lel (Erlang) jelölni
- Pl.: egy üzleti előfizető
  - $\lambda = 3$ [1/óra]
  - $h = 3$ [perc]
  - A = 3 [1/óra] · 0.05 [óra] = 0,15 [Erl]
- Pl.: központ 10 000 előfizetővel
  - $\lambda = 4000$ [1/óra]
  - $h = 3$ [perc]
  - A = 4000 [1/óra] · 0.05 [óra] = 200 [Erl]

- A t idő alatt beérkezett hívások száma λt paraméterű Poisson eloszlással adható meg.
- A beérkezések (hívások kezdete) között eltelt időt λ paraméterű exponenciális eloszlás írja le.

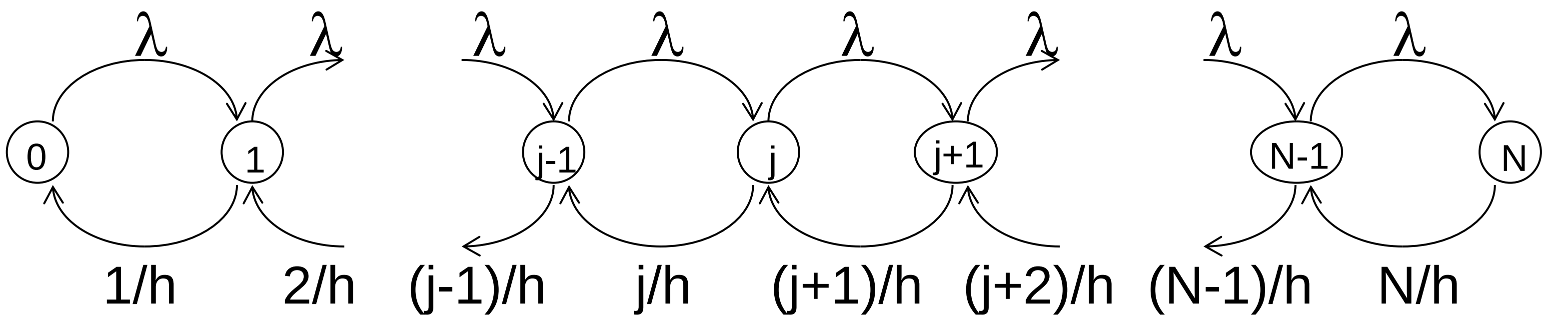
A matematikai modellekről



- A következő fóliákon lévő levezetések viszonylag egyszerűek
- Ugyanakkor többé-kevésbé feltételezik a Markov láncok ismeretét
- Ezzel együtt bent hagytam az előadásban, mert érdekes és nem bonyolult
- De *a címsorban (*)-gal megjelölt fóliákat nem kell tudni a vizsgára!*

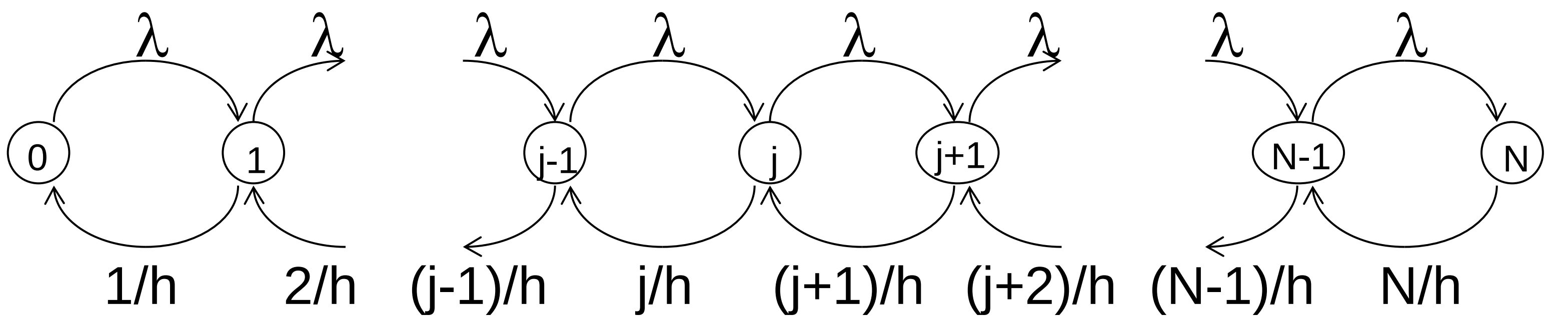
Az Erlang modell (*)

- E rendszer modellezésére használjuk az ún. folytonos idejű Markov láncot!



- Az állapotok jelentése: ennyi vonal foglalt éppen
- Pl. *j.* állapotban maradunk amíg:
  - be nem fejeződik egy hívás: *j/h* paraméterű exp. eloszlás adja meg, hogy ez mikor következik be, *j-1.* állapot a következő
  - új hívás érkezik: *λ* paraméterű exp. eloszlás adja meg, hogy ez mikor következik be, *j+1.* állapot a következő
  - abba az irányba megyünk, amelyik esemény előbb történik meg
- Markovi tulajdonság: a jövő csak a jelentől függ, a múlttól nem
  - Ha most *k* db. vonal foglalt, akkor mindegy, hogy korábban *k+1* vagy *k-1* volt foglalt
  - Az exponenciális eloszlás örökifjú tulajdonsága miatt az is mindegy, mennyi ideje vagyunk az adott állapotban

Az Erlang modell (*)



- *Fontos!* Feltesszük, hogy elég sok felhasználó van, ezért egy új igény beérkezése után is *λ* marad az előrelépési intenzitás
  - reálisan csökkennie kéne, hiszen már kevesebb felhasználó van, aki hívást kezdeményezhet
  - ez az Erlang modellnek a lényege, egyben a gyengéje
  - de ettől lesz egyszerű
  - többszáz felhasználó esetén jól használható közelítés

Az Erlang modell (*)

- Próbáljuk ennek segítségével megoldani az eredeti problémát!
- Legyen $P_i(t)$ annak a valószínűsége, hogy a rendszer az *i.* állapotban van a *t* időpontban
- Nézzük meg a rendszer állapotát egy kis Δt idővel később!
  - feltesszük, hogy addig max. egy új hívás érkezik <kizáró vagy> egy hívás ér véget
  - annak a valószínűsége, hogy egy hívás véget ér ≈ $\frac{1}{h}\Delta t$
  - annak a valószínűsége, hogy új hívás érkezik ≈ $\lambda\Delta t$
  - Δt idő múlva 0 áramkörünk foglalt ha:
    - most 1 áramkör használt és a vizsgált Δt idő alatt egy hívás befejeződik:

$$P_1(t)\cdot\left(\frac{1}{h}\Delta t\right)$$

    - vagy most is 0 áramkör foglalt és nem érkezik hívás Δt idő alatt:

$$P_0(t)\cdot(1-\lambda\Delta t)$$

Az Erlang modell (*)

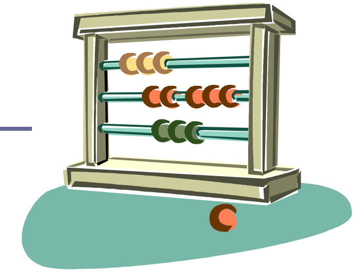
- Nézzük meg a rendszer állapotát egy kis Δt idővel később!
  - *Δt* idő múlva *j* áramkörünk foglalt ha:
    - most *j-1* van használatban és egy új hívás érkezik
    - vagy ha most *j+1* használt és egy hívás befejeződik
    - vagy ha most is *j* áramkört használunk és *Δt* idő alatt nem keletkezik, és nem fejeződik be hívás
  - *Δt* idő múlva *N* áramkörünk akkor lesz foglalt ha
    - most *N-1* használt és *Δt* idő alatt keletkezik egy hívás
    - vagy ha most is *N* használt és *Δt* idő alatt nem fejeződik be hívás
- Ezek alapján:

$$P_0(t+\Delta t) = P_0(t)\cdot(1-\lambda\Delta t) + P_1(t)\cdot\left(\frac{1}{h}\Delta t\right)$$

$$P_j(t+\Delta t) = P_{j-1}(t)\cdot\lambda\Delta t + P_j(t)\cdot\left(1-\lambda\Delta t-\frac{j}{h}\Delta t\right) + P_{j+1}(t)\cdot\left(\frac{j+1}{h}\Delta t\right)$$

$$P_N(t+\Delta t) = P_{N-1}(t)\cdot\lambda\Delta t + P_N(t)\cdot\left(1-\frac{N}{h}\Delta t\right)$$

Az Erlang modell (*)



- Nézzük most az első egyenletet:

$$P_0(t+\Delta t) = P_0(t)\cdot(1-\lambda\Delta t) + P_1(t)\cdot\left(\frac{1}{h}\Delta t\right)$$

$$P_0(t+\Delta t) = P_0(t) - P_0(t)\lambda\Delta t + P_1(t)\cdot\left(\frac{1}{h}\Delta t\right)$$

$$P_0(t+\Delta t) - P_0(t) = -P_0(t)\lambda\Delta t + P_1(t)\cdot\left(\frac{1}{h}\Delta t\right)$$

$$\frac{P_0\left(t+\Delta t\right)-P_0\left(t\right)}{\Delta t} = -P_0(t)\lambda + P_1(t)\cdot\left(\frac{1}{h}\right)$$

$$\frac{P_0\left(t+\Delta t\right)-P_0\left(t\right)}{\Delta t} = P_1(t)\cdot\left(\frac{1}{h}\right) - P_0(t)\lambda$$

$$\lim_{\Delta t\to 0}\frac{P_0\left(t+\Delta t\right)-P_0\left(t\right)}{\Delta t} = \frac{dP_0\left(t\right)}{dt}$$

$$\frac{dP_0\left(t\right)}{dt} = P_1(t)\cdot\left(\frac{1}{h}\right) - P_0(t)\lambda$$

Az Erlang modell (*)

- Itt tartunk: $$\frac{dP_0(t)}{dt} = P_1(t)\cdot\left(\frac{1}{h}\right) - P_0(t)\lambda$$
- Tudjuk azonban, hogy a legforgalmasabb órákra tervezünk, amikor:
  - a foglaltsági valószínűségek nem függenek a vizsgált időponttól (stacionárius rendszer): $P_j(t) = P_j$ (konstans)
  - így a valószínűség megváltozása: $\frac{dP_j(t)}{dt} = 0$
- Így:

$$0 = P_1\cdot\left(\frac{1}{h}\right) - P_0\lambda \quad \text{és} \quad \lambda\cdot h = A$$

$$0 = P_1 - A\cdot P_0$$

$$P_1 = A\cdot P_0$$

Az Erlang modell (*)

- Itt tartunk: $P_1 = A \cdot P_0$
- Hasonlóan a többi egyenletre ugyanezt végrehajtva kapjuk:

$$P_j = \frac{A^j}{j!} \cdot P_0$$

- Tudjuk továbbá: $\sum_{j=0}^{N} P_j = 1$
- Innen: $P_0 \cdot \sum_{j=0}^{N} \frac{A^j}{j!} = 1$ és így $P_0 = \frac{1}{\sum_{i=0}^{N} \frac{A^i}{i!}}$
- Összerakva: $P_j = \frac{\frac{A^j}{j!}}{\sum_{i=0}^{N} \frac{A^i}{i!}}$

Az Erlang modell (*)

- Itt tartunk: $$P_j = \frac{\frac{A^j}{j!}}{\sum_{i=0}^{N} \frac{A^i}{i!}}$$
- Blokkolás akkor van, ha mind az *N* vonal foglalt: $$\boxed{P_N = \frac{\frac{A^N}{N!}}{\sum_{i=0}^{N} \frac{A^i}{i!}}}$$
- Azt a legkisebb *N*-et keressük, amire P_N kisebb lesz a megengedett maximális blokkolásnál
- azonban *N* ebből zárt alakban nem fejezhető ki
- ennek ellenére konkrét számokra iteratív módszerrel viszonylag hamar meghatározható a képletből

Az Erlang modell (*)

- Tudjuk tehát, hogy az idő hányad részében foglalt minden vonal („időtorlódás”)
- Feltettük, hogy a hívások érkezési intenzitása független a fennálló hívások számától és az időtől is
- Így az összes vonal foglaltságakor is ugyanolyan valószínűséggel érkezik hívás, mint egyébként
- Emiatt a kiszámolt időtorlódás megegyezik a keresett blokkolással

Távbeszélő-hálózatok forgalmi jellemzése

- Innen: „legkevesebb hány ák. kell, hogy a blokkolás adott érték alatt maradjon?”
- *Erlang B képlete*
  - P_N -- mind az *N* vonal foglalt lesz:

$$P_N = \frac{\frac{A^N}{N!}}{\sum_{i=0}^{N} \frac{A^i}{i!}}$$

  - ez veszteséges rendszerre jó. Sorbanállásosra Erlang C -- bonyolultabb.

Agner Krarup Erlang
1878 - 1929
dán matematikus, a forgalomelmélet megalapozója
(A képlet egy 1917-es publikációjában jelent meg.)

Távbeszélő-hálózatok forgalmi jellemzése

- Pl. 1000 üzleti előfizető n vonalon:

  λ = 1000*3 [1/óra]

  h = 3 [perc]

  A = 1000*3 [1/óra]* 0.05 [óra] = 150 [Erl]

  Ekkor:

n	100	150	155	160	200
P(n)	34%	6,2%	4,3%	2,8%	0,0015%

  Nagy előfizetőszámra az elfogadható n A-hoz konvergál

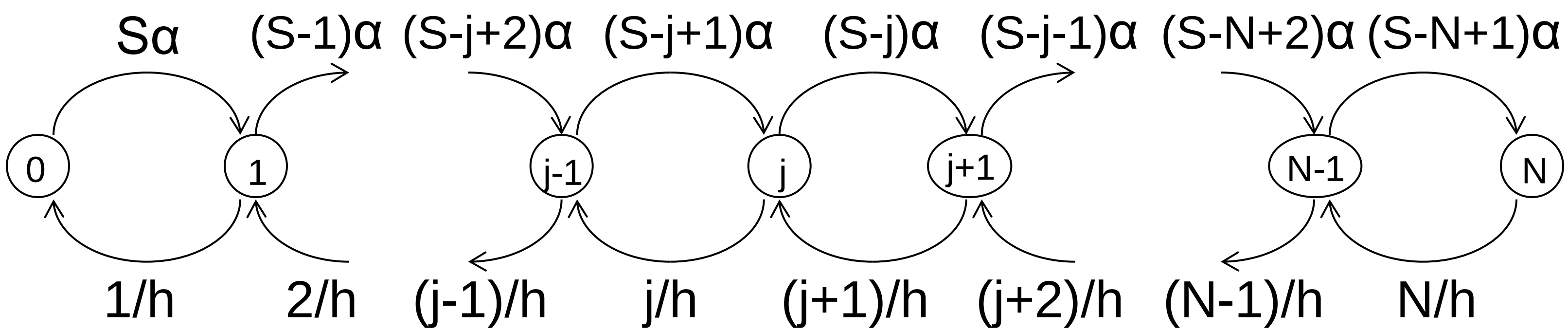
- A képletet használták pl. modem pool-ok méretezésére is

Az Engset modell

- Ha lényegesen kevesebb az előfizető (pl. kisvállalati központot kell méretezni), akkor az Erlang modell nem használható
  - Pl. a képlet szerint 3 előfizető, 3 vonal esetén sem lesz 0 a blokkolás
- Ekkor a precízebb, de bonyolultabb Engset modell alkalmazandó

Az Engset modell (*)

- Az Engset modell:



- Ahol α: egy felhasználó által generált hívások gyakorisága is exponenciális eloszlású, ennek a paramétere
- S: a felhasználók száma

Az Engset modell (*)

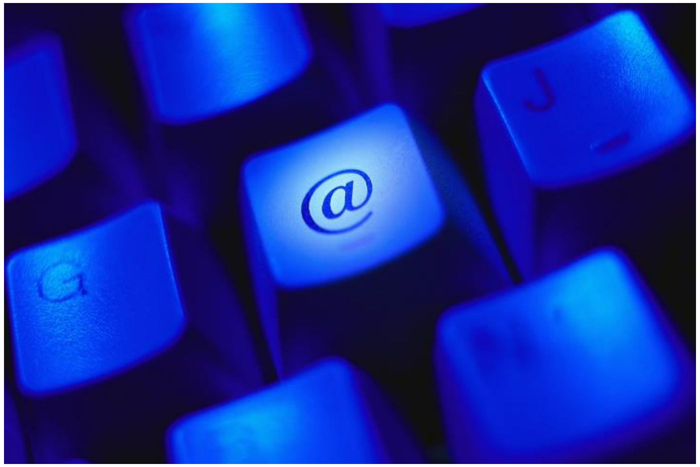
- A levezetést mellőzve a blokkolás:

$$B_{N,S} = \frac{\binom{S-1}{N} \cdot A^N}{\sum_{j=0}^{N} \binom{S-1}{j} \cdot A^j} \quad \text{ahol } A = \alpha \cdot h$$

- Itt sem tudjuk *N*-et kifejezni
- Sőt, a numerikus számítás is elég bonyolult

Forgalmi követelmények, hálózatméretezés

- Távbeszélő-hálózatok forgalmi jellemzése
- Csomagkapcsolt hálózatok forgalmi jellemzése
- Egy gondolat a teljesítőképességről



Követelmények IP hálózatokban

- Sokféle alkalmazás, sokféle követelmény
- Alkalmazások, pl.:
  - e-mail
  - telefonálás
  - videotelefonálás
  - film megnézése valós időben

IP hálózatok forgalmi modellezése

- Cél: hálózatméretezés tudományos megalapozása
- Távbeszélő-hálózatokénál lényegesen nehezebb, mert:
  - alkalmazások:
    - sokféle, különféle hálózati igényekkel
    - időben, térben változó összetételű alkalmazás-mix
    - évről évre jelentős változások lehetnek a tipikusan használt alkalmazásokban (nehéz középtávra tervezni)
    - alkalmazások erőforrásigénye is nehezen meghatározható (pl. e-mail hossza bájtban)
  - elasztikus folyamok
    - adott mennyiségű adatot kell átküldeni, de nem nagyon számít, hogy mennyi idő alatt
    - pl. FTP, HTTP, e-mail továbbítás
    - a rendelkezésre álló teljes sávszélességet elfoglalják
    - nehezen definiálható a sávszélességigény

IP hálózatok forgalmi modellezése

- Távbeszélő-hálózatokénál lényegesen nehezebb, mert: (folyt.)
  - nem független források:
    - elasztikus folyamok és a TCP garantálja a közel teljes sávszélesség kihasználást
    - emiatt blokkolás, különböző források csomagjai versengenek a továbbításért
    - követk.: nem független források
  - az egyes források leírása is bonyolultabb
- Az eddigiek következményei:
  - Hosszú távú összefüggés (időben távoli értékek is korreláltak)
  - Önhasonlóság: különböző időskálákon nézve is hasonló forgalmi jelleg (forgalom: bit/s, csomag/s) – fraktálszerű képet mutat
  - Nagy börsztösség, csomósodás
  - PSTN: n-szeres felhasználó, forgalom átlaga is n-szeres, de szórása $\sqrt{n}$-szeres: a forgalom „kisimul”
  - TCP/IP: a forgalom sokkal lassabban „simul ki”

IP hálózatok forgalmi modellezése

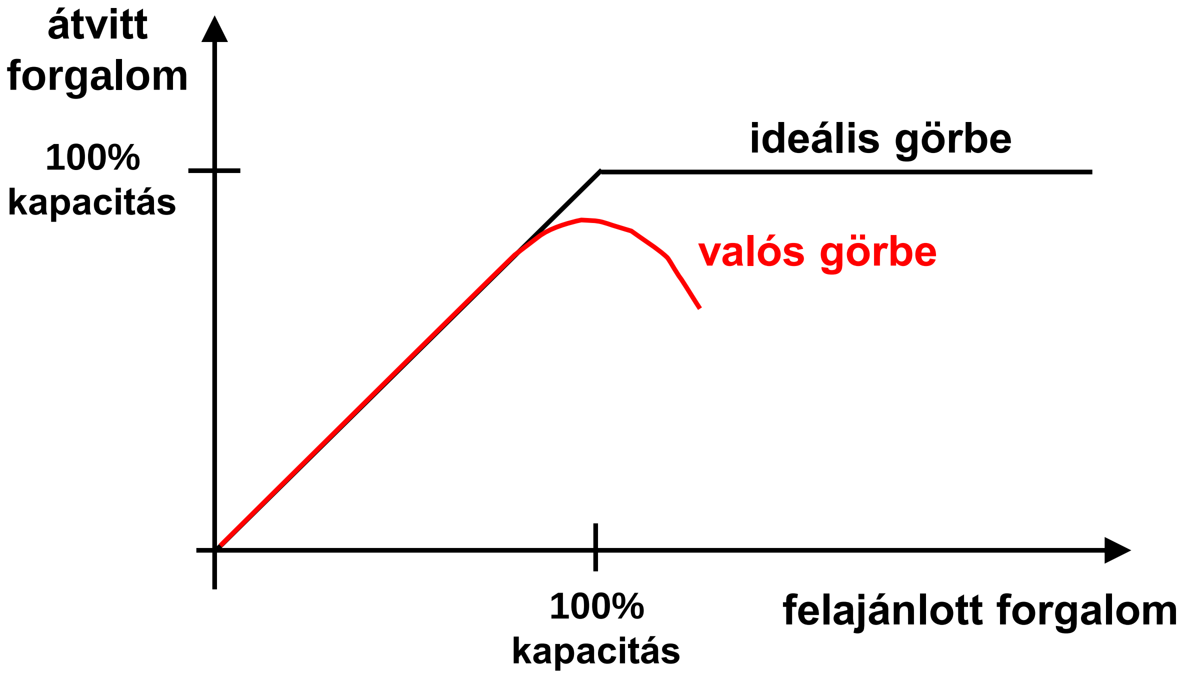
- Ezek miatt a TCP/IP forgalommodellezés még gyerekcipőben jár
  - bár vannak bíztató eredmények
- Akkor hogyan lehet TCP/IP hálózatot méretezni?
  - tapasztalatok alapján
  - mérések alapján
  - túlméretezés (overprovisioning)
    - másik ok a túlméretezés mellett: olcsó a kapacitás, de jelentős a bevétel: nem szabad egy vevőt sem elszalasztani kapacitáshiány miatt

Forgalmi követelmények, hálózatméretezés

- Távbeszélő-hálózatok forgalmi jellemzése
- Csomagkapcsolt hálózatok forgalmi jellemzése
- Egy gondolat a teljesítőképességről

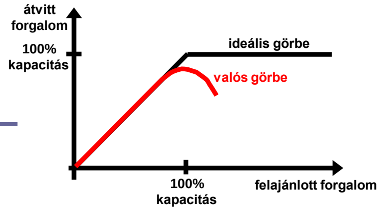
Hálózatok teljesítőképessége

- Egy meglepő megfigyelés:



- Ok: hálózat túlterhelődik
- Ennek elkerülésére célszerű a hálózatot a maximális átvitel munkapontjában megtartani

Hálózatok teljesítőképessége



Lehetséges megoldások:

- Torlódásmenedzsment
  - A kialakult torlódást kezelni, hogy ne legyen nagyobb
    - Pl. a források vegyenek vissza az adási sebességből (TCP)
- Torlódáselkerülés
  - Még a torlódás kialakulása előtt beavatkozni
  - Pl. hívásengedélyezés (Call Admission Control, CAC): csak akkor engedünk be egy hívást a hálózatba, ha az nem fogja túlterhelni azt
    - Pl. egy közértben csak adott számú kosár van, belépés csak kosárral
    - H.323 hálózatban is van erre lehetőség
    - Pl. a Rapidshare is tud ilyet
    - Pl. a Neptun is tud ilyet