

5. gyakorlat – Teljesítménymodellezés

Elméleti összefoglaló

Dimenzióanalízis. A teljesítménymodellezés feladatok megoldása során érdemes a fizikából ismert dimenzióanalízist¹ elvégezni. Idézzük fel a négyzetes úttörvény: $s = v_0 t + \frac{a}{2} t^2$

Dimenziókkal: $s[\text{m}] = v_0 \left[\frac{\text{m}}{\text{s}} \right] t[\text{s}] + \frac{a}{2} \left[\frac{\text{m}}{\text{s}^2} \right] t^2[\text{s}^2] = v_0 t[\text{m}] + \frac{a}{2} t^2[\text{m}]$

A dimenzióhasználat fő motivációja, hogy ha a dimenziók nem stimmelnek, akkor a képletet is biztosan elrontottuk valahol.² A dimenzióanalízis gyakran segít a megfelelő képlet kiválasztásában. Fontos, hogy a „darab”, „kérés” stb. jellegű mértékegységek nem számítanak külön dimenziónak, ezért pl. a $\frac{\text{kérés}}{\text{s}}$ és az $\frac{1}{\text{s}}$ dimenziók megegyeznek.

Alapképletek. Little-törvény: $N = X \cdot T$, $N [1] = X \left[\frac{1}{\text{s}} \right] \cdot T [\text{s}]$

Kihasznátság intuitíven és a Little-törvényből *egyetlen kizárólagos* erőforráspéldány esetén:

$$U = \frac{X}{X_{\max}} = \frac{T_{\text{busy}}}{T_{\text{measured}}} = N = X \cdot T$$

Átbocsátóképesség végrehajtási időből *egyetlen kizárólagos* erőforráspéldány esetén (az átbocsátóképesség az elérhető legnagyobb átbocsátás, vagyis ilyenkor a kihasznátság 100%): $X = \frac{U}{T} \Rightarrow X_{\max} = \frac{1}{T}$

1. Diszk teljesítménye

Egy diszk 50 kérést szolgál ki másodpercenként. Minden kérés kiszolgálása 0,005 másodpercet vesz igénybe. A rendszerben nincs átlapolódás.

- Mekkora a maximálisan kiszolgálható terhelés (érkezési ráta)?
- Mekkora a kihasznátság?

2. Zárthelyi megtekintése

A zárthelyik megtekintése során a hallgatónak lehetőségük van reklamálni esetleges javítási hibák miatt. Sikeres reklamáció esetén a pontszámuk módosításra kerül. Az első nagyfeladatból (F1) óránként 10 darabot képes átnézni egy javító, a második nagyfeladatból (F2) pedig 20 darabot. Mindkét feladathoz tartozik 1-1 javító, akik az adott feladatot javították. A továbbiakban készítsünk minden kérdéshez egy-egy folyamatmodellt és határozzuk meg, hogy óránként hány hallgató dolgozatát sikerül átnézni az egyes esetekben!

- A hallgatók először az F1, majd az F2 feladatot nézetik át a javítóval.
- A leleményes hallgatók a két feladatot külön-külön már egyszerre két javítónak adják oda, mivel külön lapra voltak írva. Mit nyerünk a párhuzamosítással?
- A nagy tömeg miatt a hallgatók csak az egyik feladatukat nézetik át, mégpedig azt, amelyiknek a javítója éppen szabad.
- Híre ment, hogy a második feladat javítója sokkal kevésbé szigorú, így a hallgatók 80%-a inkább kivárja ennél a javítónál a sort. A maradék 20% a másik javítónál reklamál az első feladattal kapcsolatban.
- A hallgatók 10%-ának a reklamáció után már csak 1-2 pont kellene a jobb jegyhez, ezért újra és újra megpróbálkoznak a reklamációval. Feltételezhetjük, hogy a hallgatók az a) részben leírt reklamációs stratégiát használják.
- Mi történne másként, ha bármelyik javító bármelyik feladatot hajlandó átnézni (de az egyes feladatok átnézése változatlan ideig tart), és így szeretnék a hallgatók az eredeti tervnek megfelelően először az F1, majd az F2 feladatot átnézni a javítóval?

3. Kétrétegű architektúra

Adott egy webszerver (WS) és két fürtözött adatbázisszerver (DB1, DB2). A két adatbázis szerver közt súlyozott round robin terheléelosztás alapján választunk, 1:2 arányban. Minden felhasználói kérés kiszolgálása során mindkét fajta erőforrást használjuk. A csúcsideőszakban 30 percig monitorozzuk a rendszert, ezalatt 9000 kérést szolgál ki. A szerveken mért foglaltsági idők: WS – 1350 s CPU idő; DB1 – 810 s, DB2 – 1320 s diszk IO idő.

- Készítsünk folyamatmodellt a kérések feldolgozásáról a szöveg alapján!
- Mekkora az egyes szerverek jelenlegi átbocsátása?
- Mennyi időt töltenek egy-egy hozzájuk beérkezett kérés kiszolgálásával a szerverek?

¹Dimenzióanalízis (Wikipédia), <http://hu.wikipedia.org/wiki/Dimenzió>

²Ajánlott olvasmány: what if? – Droppings, <http://what-if.xkcd.com/11/>

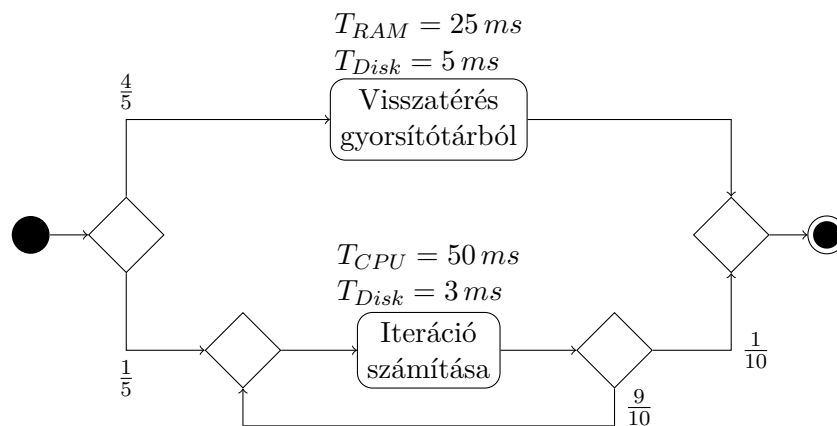
- d) Mekkora a rendszer maximális áteresztőképessége?
- e) Miért nem egyféle foglaltsági időt vettünk figyelembe a két erőforrástípusnál?
- f) Hol csal még így is a modell?

4. Mikroszolgáltatás (* korábbi zárthelyi feladat)

Mikroszolgáltatásunk egy közelítő algoritmust implementál, melyet más szolgáltatások hálózaton keresztül építőkockaként felhasználhatnak a saját funkciójuk megvalósításához. Mivel sok egyforma kérés érkezik, a szolgáltatás gyorsítótárazza a válaszokat, így a korábban kiszámolt eredmények gyorsan visszaküldhetők. Maga a számítás egy igen költséges iteratív folyamat, amelyet addig kell ismételni, amíg a megoldás kellően precíz nem lesz.

Az alábbi modell a kísérleti futtatás tapasztalatait összegzi. Megfigyeléseink szerint a kérések 80%-a gyorsítótárból kiszolgálható, míg egy iteratív számítási lépés végén átlagosan 10% eséllyel lesz elég precíz a megoldás. A kétféle tevékenységhez megmértük az egyes erőforrások átlagos foglaltsági idejét is, ezt az alábbi ábrán a lépések felett tüntettük fel (csak a nem elhanyagolható ideig használt erőforrásokra).

A méréshez az élesben is használt szerveret használtuk, amely 2 CPU-val, 1 RAM modullal és 1 diszkkal rendelkezik.



- a) Ha mindegyik lépés esetén csak a leghosszabban használt erőforrást vesszük figyelembe, mekkora lehet a rendszer átbocsátóképessége?
- b) Ha csak a mindkét lépés által jelentősebben használt, egyetlen közös diszk erőforrást tekintjük, akkor mekkora lehet a rendszer átbocsátóképessége?
- c) Mekkora ez alapján a rendszer tényleges átbocsátóképessége? Melyik erőforrás (fel)skálázásával lehetne még tovább növelni?
- d) Ha a kihasználtság 50%-os, és egy kérés beérkezésétől az eredmény visszaküldéséig (várakozással együtt) átlagosan 100 ms telik el, akkor átlagosan hány olyan kérés van a rendszerben, amelynek a feldolgozása még nem fejeződött be?

5. Sziget közlekedési hálózata (* korábbi zárthelyi feladat)

Egy sziget lakói minden reggel munkába menet átkelnek a szigetet ölelő tavon. Észak felé híd vezet, dél felé autósomp. Az irányonként egysávos híd 200 m hosszú, és 60 km/h sebességgel szabad rajta haladni, a követési távolság (hátsó lámpától hátsó lámpáig 30 m) betartása mellett. A négy komphajó egyenként 15 percenként teszi meg a sziget-szárazföld-sziget kört, és így óránként négyen együtt legfeljebb 800 autót tudnak átvinni a szárazföldre.

- a) Mekkora a híd átbocsátóképessége (észak felé)?
- b) Hány autó fér el egy komphajón?
- c) A reggeli csúcsgorgalomban mekkora a szigetet elhagyó két útvonal együttes átbocsátóképessége?
- d) Ha délben a szárazföldi főutat baleset miatt lezárták, és a szigeten keresztül (a hídon, majd a kompon átkelve) terelik a forgalmat, mekkora a terelőútvonal átbocsátóképessége?
- e) Valamelyik reggel 7:00 és 8:30 között 900 autó hagyta el a szigetet komppal. Mennyi volt ebben az időszakban a kompok átbocsátása és kihasználtsága?
- f) A fenti mérésben átlagosan hány autó állt sorba egyszerre a parton, ha az autók jól időzítve, átlagosan fél perccel a beszállásuk előtt érkeztek kompkikötőhöz?

6. Tudásbázis (*)

Vállalatunk nyilvános szakmai tudástára egymásra is hivatkozó szócikkeket kínál a cég termékeit világszerte használó ügyfeleknek. Egyetlen szócikk lekérésének kiszolgálásához a szervert átlagosan 60 ms-ig veszi igénybe. A szócikk megtekintése után az olvasó csak az esetek 30%-ában hagyja el az oldalt, többnyire ugyanis egy újabb szócikkre mutató hivatkozásra kattint.

- a) Egy olvasó összes tudásszomjának kielégítéséhez átlagosan mekkora szerveridő szükséges?
- b) Tekintsük úgy, hogy az egyes kérések a szerveren nem párhuzamosíthatóak. Óránként hány egyedi látogatót képes kiszolgálni a szerver?