PROCESSZÁLÁSI IDŐ SKÁLÁZÁSA VS. LÁTENCIA

Tétel: Ha:

- az új processzállás K-szor gyorsabb
- a bejövő kérések frekvenciája független (!) a processzálás idejétől

akkor:

- az átlagos látencia visszaesik legalább a korábbi látencia K-adrészére (nem lesz nagyobb ennél).

Bizonyítás vázlat

Legyen

- az i. beérkezési idő (az előzőhöz képest): t_{IN}^{i} speciálisan i = 1-re az értéke legyen 0
- az i.e processzálási idő: t_{PROC}^{i}
- az új processzálási idő a p-ed része a korábbinak (p<1)

Látencia az eredeti esetben: Az i. látencia az eredeti processzálás esetén a következő:

$$(1)\sum_{i=1}^{L} t_{PROC}^{i} - \sum_{i=1}^{L} t_{IN}^{i}$$

Mindaddig ($i<=l_0$) amíg a processzálás lasabb mint az új beérkezés:

$$(2) L. processzálás = \sum_{i=1}^{L} t_{PROC}^{i} \geq \sum_{i=1}^{L+1} t_{IN}^{i} = L+1. beérkezés$$

Utána „újrakezdődik" a számlálás…

Így is el lehet mondani: Az (1, 2) azt fejezi ki, hogy a látenciák összeadódnak, mindaddig, amíg gyorsabban érkeznek be a kérések, mintahogy a rendszer processzálja őket. Amikor viszont elfogynak a kérések, akkor elölről kezdődik a számolás, a korábbi látenciákat el lehet „felejteni", nincsenek hatással az új látenciákra.

Látencia az új esetben: Az új processzálás esetén az i. látencia a következő:

$$(3) p * \sum_{i=1}^{L} t_{PROC}^{i} - \sum_{i=1}^{L} t_{IN}^{i}$$

Mindaddíg ($i <= I_{1,1}$), amíg a processzálás lasabb mint az új beérkezés.

Belátható, hogy $I_{1,1} <= I_0$

$i > I_{1,1}$ esetén az új látencia így alakul:

$$(4) p * \sum_{i=I_{1,1}+1}^{L} t_{PROC}^{i} - \sum_{i=I_{1,1}+1}^{L} t_{IN}^{i}$$

Mindaddíg ($i <= I_{1,2}$), amíg a processzálás megint lasabb mint az új beérkezés. Belátható, hogy a fenti legfeljebb akkora, mint a (3) alatti:

$$(5) p * \sum_{i=I_{1,1}+1}^{L} t_{PROC}^{i} - \sum_{i=I_{1,1}+1}^{L} t_{IN}^{i} < p * \sum_{i=1}^{L} t_{PROC}^{i} - \sum_{i=1}^{L} t_{IN}^{i}$$

S.í.t általában belátható, hogy az új processzálás esetén a látencia értéke legfeljebb a (3) alatti.

$$(6) \text{új } i.\,látencia \ \leq p * \sum_{i=1}^{L} t_{PROC}^{i} - \sum_{i=1}^{L} t_{IN}^{i}$$

Látenciák összehasonlítása

A tétel állítása (6)-ból és az alábbiból következik:

$$(7) p < 1 \Rightarrow p * \sum_{i=1}^{L} t_{PROC}^{i} - \sum_{i=1}^{L} t_{IN}^{i} < p * \left[\sum_{i=1}^{L} t_{PROC}^{i} - \sum_{i=1}^{L} t_{IN}^{i}\right]$$

Szavakkal: minden új látencia, legfeljebb a p-szerese a korábbinak. Így akkor az átlag is legfeljebb a p-szerese.

További kérdések

- Mi van a **percentile**-kkel? Beláthatónak tűnik, hogy a fenti a percentile-kre is igaz. Ez is abból következik, hogy minden új látencia legfeljebb a p-szerese a korábbinak. Így belátható, hogy a 99%-os értéknél is ez lesz a helyzet (akkor is, ha helycsere van!)

- Mi van, ha nem a proci gyorsabb, hanem **párhuzamosítunk** (Amdhal és Universal Scalability)?
- Mi van, ha a bejövő **kérések frekvenciája nem független** a feldolgozás sebességétől/látenciától? Ami egy izgalmas kérdésnek tűnik
- Mivel lehetne (szemléletesen/jól) **jellemezni** az összefüggést a bejövő kérések frekvenciája, a processzálási idő és látencia között?