Laborator 2

Evaluarea performantelor calculatoarelor paralele

Probleme

1. S-a analizat factorul de accelerare (*S*) pentru un sistem cu *n* procesoare in raport cu un sistem uniprocesor, in conditiile in care $f_i = 1/n$ este probabilitatea de a asigna aceeasi problema la *i* procesoare cu o incarcare medie $d_i = 1/i$, pentru $i = 1, 2, 3, \ldots , n$.

a) Sa se refaca calculul accelerarii in cazul unei distributii noi de probabilitate a modurilor de operare:

$$f_i = \frac{i}{\sum_{i=1}^{n} i}$$

pentru i = 1, 2, 3, ... , n.

b) Analog, pentru o alta distributie de probabilitate:

$$f_i = \frac{n-i-1}{\sum_{i=1}^{n} i}$$

pentru $i = 1, 2, 3, \ldots , n$.

Solutie

a) Timpul necesar pe sistemul paralel este:

$$T_n = \sum_{i=1}^{n} f_i d_i = \sum_{i=1}^{n} \left(\frac{i}{\sum_{i=1}^{n} i} \cdot \frac{1}{i} \right) = \sum_{i=1}^{n} \frac{1}{\frac{n(n+1)}{2}} = \frac{2}{n(n+1)} \cdot n = \frac{2}{n+1}$$

Factorul de accelerare este:

$$S = \frac{T_1}{T_n} = \frac{n+1}{2}$$

b)

$$T_n = \sum_{i=1}^{n} f_i d_i = \sum_{i=1}^{n} \left(\frac{n-i-1}{\sum_{i=1}^{n} i} \cdot \frac{1}{i} \right) = \frac{2}{n(n+1)} \left[(n-1)\sum_{i=1}^{n} \frac{1}{i} - \sum_{i=1}^{n} 1 \right] = \ldots$$

Factorul de accelerare este:

$$S = \frac{T_1}{T_n} = \ldots \le \frac{n(n+1)}{2n(\ln n - 1) - 2\ln n}$$

2. Un sistem multiprocesor este capabil de o rata maxima de executie de 100 MFLOPS. Codul scalar este prelucrat cu o rata de 1 MFLOPS. Care este performanta masinii daca 10% din cod este secvential si 90% este paralel ?

Solutie

Deoarece se considera numai doua rate de executie (secvential si paralel) se va utiliza legea lui Amdahl, in care:

$R_H = 100$

$R_L = 1$

$f = 0.9$

$1-f = 0.1$

$$R = \frac{1}{\frac{f}{R_H} + \frac{1-f}{R_L}} = \frac{1}{\frac{0.9}{100} + \frac{0.1}{1}} = 9.17\ MFLOPS$$

3. O cale de extindere a legii lui Amdahl este sa se studieze R ca o functie de R_H/R_L, cand f este definit pe un interval $f\in [a,b]$, cu $0<a<b<1$.

a) Sa se deduca expresia pentru rata medie si sa se interpreteze rezultatul.

b) Se poate observa o legatura cu supozitia lui Minsky?

Solutie

a) Se noteaza $R_{HL} = R_H / R_L$. Atunci rata medie este:

$$R_{med} = \frac{1}{b-a}\int_a^b R(f)df = \frac{1}{b-a}\int_a^b \frac{R_H df}{f(1-R_{HL}) + R_{HL}}$$

$$R_{med} = \frac{R_H}{(b-a)(1-R_{HL})}\ln\frac{b(1-R_{HL}) + R_{HL}}{a(1-R_{HL}) + R_{HL}}$$

Aceasta functie este aproximativ log x/x. Pentru $a = 0$ si $b = 1$ se obtine:

$$R_{med} = R_H\,\frac{\ln R_{HL}}{R_{HL} - 1}$$

Rezulta ca $R_{med} \to 0$ cand R_{HL} devine foarte mare.

b) Supozitia lui Minsky arata ca performanta unui sistem paralel creste cel mult ca o functie logaritmica de numarul de procesoare ($\log_2 n$). Din expresia de mai sus se obtine:

$$\frac{R_{med}}{R_L} = \frac{R_H}{R_L}\cdot\frac{\ln R_{HL}}{R_{HL}-1} = \frac{R_H}{R_L}\cdot\frac{\ln R_{HL}}{\frac{R_H - R_L}{R_L}}$$

Dar, $R_H - R_L \approx R_H$, deci se obtine:

$$\frac{R_{med}}{R_L} \approx \ln R_{HL}$$

Astfel, performanta sistemului paralel este limitata la o crestere logaritmica, pe masura ce R_{HL} creste.

4. Un algoritm secvential pentru descompunerea unei matrici patrate de $n*n$ elemente, in matricile triunghiulare superioara si inferioara necesita ($n^3/3 - n^2/3$) unitati de timp. Intr-un sistem paralel cu p procesoare timpul necesar pentru executia algoritmului paralel este $(n^2 - 1)(n/2)/p$.

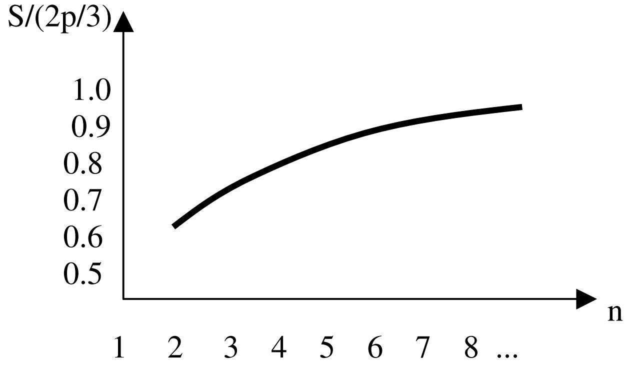
a) Sa se calculeze factorul de accelerare. Sa se reprezinte grafic in functie de n.

b) Care este eficienta acestui algoritm cu $n \to \infty$? Pentru ce valoare a lui n algoritmul este in interiorul a 5% din eficienta maxima?

Solutie

a) Factorul de accelerare pentru acest algoritm este:

$$S = \frac{T_1}{T_p} = \frac{\frac{n^3}{3} - \frac{n^2}{3}}{(n^2-1)\frac{n}{2}\cdot\frac{1}{p}} = \frac{2pn}{3(n+1)}$$



b) Eficienta:

$$E = \frac{S}{p} = \frac{\frac{2p}{3}\cdot\frac{n}{n+1}}{p} = \frac{2}{3}\cdot\frac{n}{n+1}$$

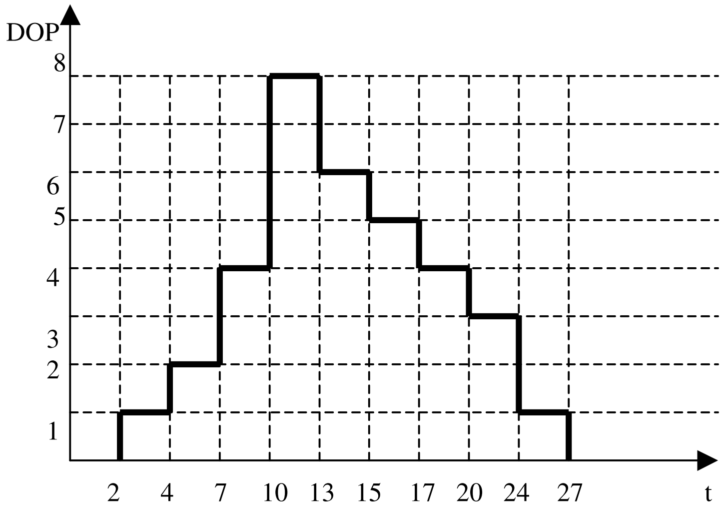
La limita se obtine:

$$\lim_{n\to\infty} E = \frac{2}{3}$$

Pentru determinarea valorii lui n se pune conditia:

$$\frac{n}{n+1} = \frac{95}{100} \quad \Rightarrow \quad n = 19$$

5, Sa se determine paralelismul mediu pentru un algoritm de tip divide-and conquer avand urmatorul profil al paralelismului.



Solutie

$A = \sum DOP_i * t_i / \sum t_i$ = (1x5+2x3+3x4+4x6+5x2+6x2+8x3)/(5+3+4+6+2+2+3)=93/25=3.72

6. Factorul de accelerare medie armonica pentru un multiprocesor (Hwang si Briggs). Se presupune $T_i = 1/i$ pentru $i = 1, 2, ..., n$. Corespunde cazului ideal cand un job unitate de timp este executat de i procesoare in timp minim. Se mai poate interpreta aceasta presupunere ca $R_i = i$, caci rata de executie creste de i ori de la $R_1 = 1$ atunci cand sunt utilizate i procesoare.

Se considera trei distributii de probabilitate in care

$$s = \sum_{i=1}^{n} i$$

si anume:

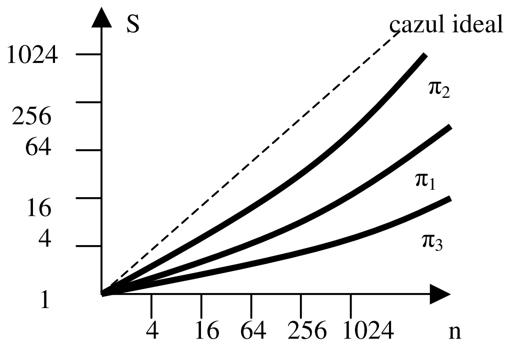
$\pi_1 = (1/n, 1/n, ..., 1/n)$ distributie uniforma

$\pi_2 = (1/s, 2/s, ..., n/s)$ favorizeaza utilizarea mai multor procesoare

$\pi_3 = (n/s, (n-1)/s, ..., 1/s)$ favorizeaza utilizarea a mai putine procesoare.

Solutie

Se reprezinta grafic $S_h^* = \dfrac{T_1}{T^*} = \dfrac{1}{\sum_{i=1}^{n} \frac{f_i}{R_i}}$ in functie de n (numar de procesoare).



Din aceste grafice rezulta ca π_2 produce o accelerare mai mare decat π_1, care la randul sau produce o accelerare mai mare decat π_3.

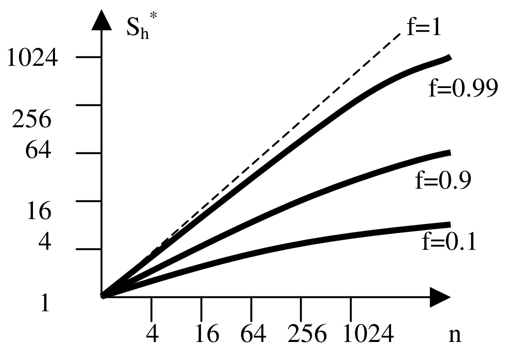
7. Se presupune $R_i = i$, cu o distributie $\pi = (1\text{-}f, 0, 0, ..., 0, f)$. Aceasta inseamna ca sistemul este utilizat fie in mod uniprocesor cu probabilitate *1-f*, fie in mod complet paralel cu probabilitatea *f*. Sa se reprezinte graficul lui S_h^* pentru diferite valori ale lui *f:* accelerarea ideala se obtine pentru *f* = *1* (complet paralel).

Solutie

Se reprezinta grafic functia:

$$S_h^* = \frac{1}{\dfrac{1-f}{R_1} + \dfrac{f}{R_n}} = \frac{1}{\dfrac{1-f}{1} + \dfrac{f}{n}} = \frac{n}{n(1-f)+f}$$

in functie de *n* (numarul de procesoare), pentru diferite valori ale lui *f*.



Daca *f* scade, atunci S_h^* scade dramatic.

8. Se considera o incarcare $O(1)=T(1)=n^3$, $O(n)=n^3+n^2\log_2 n$, $T(n)=4n^3/(n+3)$, unde n = numar de procesoare. Sa se determine accelerarea, eficienta, redundanta, utilizarea si factorul de calitate. Apoi sa se reprezinte grafic.

Solutie

Substituind in relatiile:

$$S(n)=\frac{T(1)}{T(n)}$$

$$E(n)=\frac{S(n)}{n}=\frac{T(1)}{n\cdot T(n)}$$

$$R(n)=\frac{O(n)}{O(1)}$$

$$U(n)=R(n)\cdot E(n)=\frac{O(n)}{n\cdot T(n)}$$

$$Q(n)=\frac{S(n)\cdot E(n)}{R(n)}=\frac{T^3(1)}{n\cdot T^2(n)\cdot O(n)}$$

rezulta:

$$S(n)=\frac{n+3}{4}$$

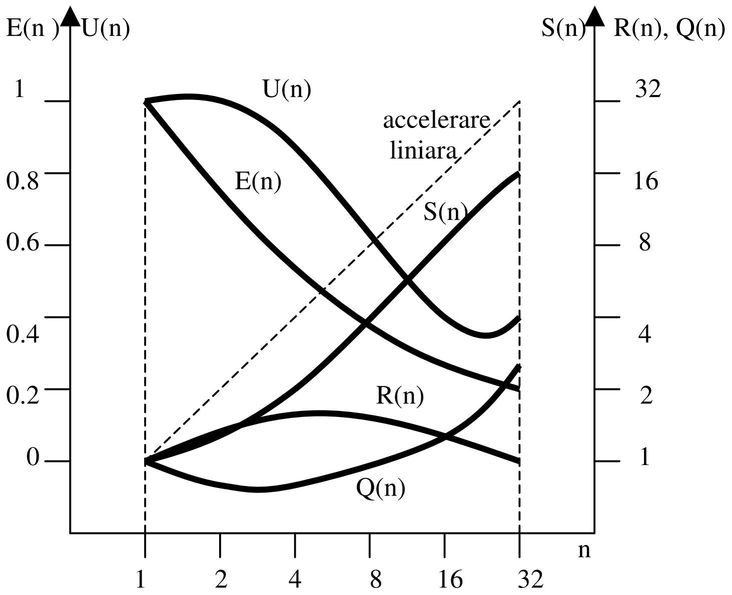
$$E(n)=\frac{n+3}{4\cdot n}$$

`

$$R(n)=\frac{n+\log_2 n}{n}$$

$$U(n)=\frac{(n+3)(n+\log_2 n)}{4\cdot n^2}$$

$$Q(n)=\frac{(n+3)^2}{16(n+\log_2 n)}$$



Se observa relatiile: *1/n ≤E(n) ≤U(n) ≤1* si *0 ≤Q(n) ≤S(n) ≤n*. Accelerarea ideala corespunde cazului ideal cu eficienta 100%.

Tema

Sa se scrie un program care sa calculeze principalele marimi ce caracterizeaza performantele unui sistem paralel (factorul de accelerare, rata de executie, eficienta, utilizarea, calitatea paralelismului). Sa se reprezinte grafic aceste marimi in functie de numarul de procesoare si dimensiunea problemei de rezolvat, avand in vedere diferite distributii de probabilitate a modurilor de lucru.