1

dalla prima lezione

Obiettivi del calcolo parallelo:

a) Risolvere uno stesso problema in minor tempo
b) Risolvere un problema piu' grande nello stesso tempo

Esempio: previsioni meteorologiche

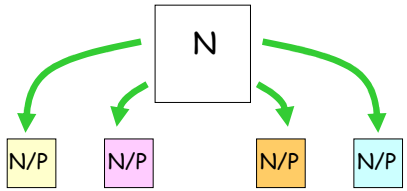
- Previsione a 3 giorni in un'ora
- Con un calcolatore di potenza doppia si puo' fare
  a) Previsione a 3 giorni in mezz'ora
  b) Previsione a 6 giorni in un'ora

2

a) uno stesso problema in minor tempo

Idea di base:

Decomposizione di un problema di dimensione N in P sottoproblemi indipendenti di uguale dimensione N/P da risolvere concorrentemente



3

misurare l'efficacia della suddivisione

Misuriamo di quanto si riduce il tempo di esecuzione

Sia T(N,1) il tempo per la risoluzione del problema di dimensione N con P=1 unita' di calcolo

Idealmente, fissato P>1, ci si aspetta che il tempo di risoluzione si riduca di P volte, cioe'

$$T(N,P) = \frac{T(N,1)}{P}$$

Alla base di tale aspettativa ci sono alcune semplificazioni:

- Suddivisione esatta del problema in P parti
- Assenza di costi dovuti alla suddivisione
- Assenza di influenza di P sul tempo di esecuzione

4

speed-up

Idealmente si ha

$$\frac{T(N,1)}{T(N,P)} = P$$

Il rapporto

$$S_P = \frac{T(N,1)}{T(N,P)}$$ **Speed Up**

misura quanto siamo vicini al caso ideale.

Ovviamente si ha:

nel caso ideale $S_P *= P$

in pratica $S_P < P$

efficienza

Il valore ottenuto dello SpeedUp potrebbe avere poco senso se non lo si confronta al valore ideale

Il rapporto

$$E_P = \frac{S_P}{S_P'} = \frac{S_P}{P}$$ **Efficienza**

Misura, in percentuale, quanto siamo vicini allo SpeedUp ideale

Ovviamente si ha

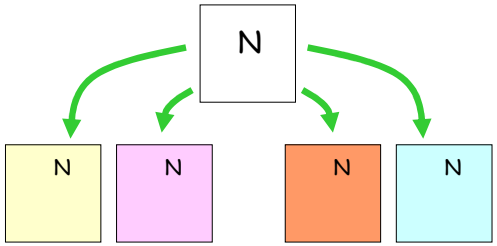
Caso ideale $E_P *= 1$

in pratica $E_P < 1$

b) un problema piu' grande nello stesso tempo

Idea di base:

Risoluzione concorrente di P problemi di dimensione N



N.B. complessivamente si risolve un problema di dimensione PN

come misurare l'efficacia della suddivisione

Misuriamo se, aumentando la dimensione dle problema, il tempo rimane costante

Sia T(N,1) il tempo per la risoluzione di un problema di dimensione N con P=1 esecutore

Idealmente, fissato P>1, ci si aspetta che il tempo di risoluzione di ciascuno dei P sottoproblemi sia

$$T(PN,P) = T(N,1)$$

Alla base di tale aspettativa ci sono alcune semplificazioni:

- Suddivisione esatta del problema in P parti
- Assenza di costi dovuti alla suddivisione
- Assenza di influenza di P sul tempo di esecuzione

efficienza scalata

Idealmente si ha

$$\frac{T(N,1)}{T(PN,P)} = 1$$

Il rapporto

$$SE_P = \frac{T(N,1)}{T(PN,P)}$$ **Efficienza Scalata**

Misura, in percentuale, quanto siamo vicini al caso ideale.

Ovviamente si ha:

nel caso ideale $SE_P{*} = 1$

in pratica $SE_P < 1$

speed up scalato

Anche il valore ottenuto dell'efficienza scalata potrebbe avere poco senso se non lo si confronta con P

Il rapporto

$$SS_P = P\ SE_P = \frac{P\ T(N,1)}{T(PN,P)}$$ **Speed Up Scalato**

Misura come usiamo i P processori

Ovviamente si ha

Caso ideale $SS_P{*} = P$

in pratica $SS_P < P$

caso ideale vs caso reale

Quali sono i fattori che impediscono di raggiungere il caso ideale?

- comunicazione / sincronizzazione
- I/O, accesso al file system
- generazione di thread e/o processi
- Inizializzazione di variabili
- Overhead per la suddivisione del problema

In generale $T(N,1)$ si puo' decomporre

$$T(N,1) = T_{seq}(N,1) + T_{par}(N,1)$$

con $T_{par}(N,1)$ (sezione perfettamente parallelizzabile)

$T_{seq}(N,1) > 0$ (sezione sequenziale)

qual e' l'impatto?

Da $T(N,1) = T_{seq}(N,1) + T_{par}(N,1)$ si ha

$$T(N,P) = T_{seq}(N,1) + \frac{T_{par}(N,1)}{P}$$

Definizione

$$\alpha = \frac{T_{seq}(N,1)}{T(N,1)}$$

e' la frazione seriale, cioe' la percentuale di codice che non e' parallelizzabile

SpeedUp

$$S_P = \frac{T(N,1)}{T(N,P)} = \frac{T(N,1)}{T_{seq}(N,1) + T_{par}(N,1)/P} = \frac{1}{\alpha + (1-\alpha)/P}$$

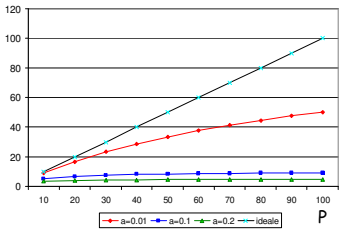
Legge di Amdhal

impatto della parte seriale

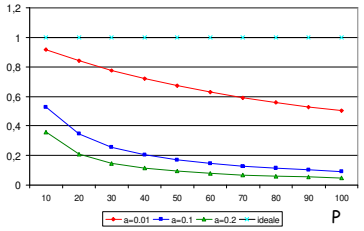
Cosa accade quando P aumenta?

$$\lim_{P\to\infty} S_P = \lim_{P\to\infty} \frac{1}{\alpha+(1-\alpha)/P} = \frac{1}{\alpha}$$

$$\lim_{P\to\infty} E_P = \lim_{P\to\infty} \frac{S_P}{P} = 0$$



Speed Up



Efficienza

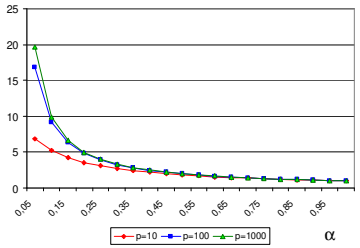
13

impatto della parte seriale

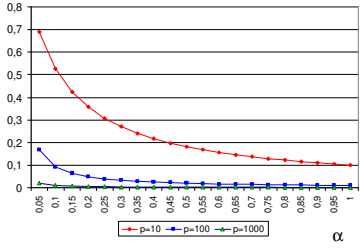
Cosa accade quando α aumenta?

$$\lim_{\alpha\to 1} S_P = \lim_{\alpha\to 1} \frac{1}{\alpha+(1-\alpha)/P} = 1$$

$$\lim_{\alpha\to 1} E_P = \lim_{\alpha\to 1} \frac{S_P}{P} = \frac{1}{P}$$



Speed Up



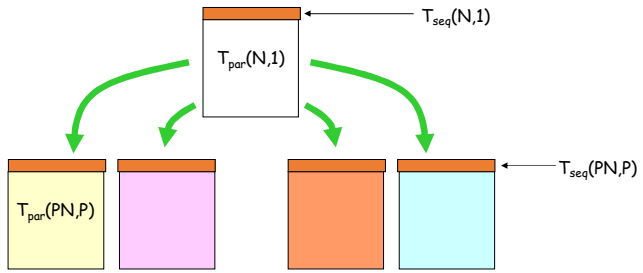
Efficienza

14

impatto della parte seriale

Che accade quando P e N aumentano tale che

1. $T_{par}(N,1) = T_{par}(PN,P) = cost$
   - (dimensione dei sottoproblemi costante al cresere di P)
2. $T_{seq}(N,1) = T_{seq}(PN,1) = cost$
   - (la parte sequenziale e' indipendente da P)



15

impatto della parte seriale

Si ha

- Dalla 1. ➔ $T_{par}(NP,1) = PT_{par}(N,1)$
- Dalla 2. ➔ $\lim_{N\to\infty} \alpha = \lim_{N\to\infty} \frac{T_{seq}(N,1)}{T(N,1)} = \lim_{N\to\infty} \frac{cost}{T(N,1)} = 0$

E quindi

$$S_P = \frac{T(PN,1)}{T(PN,P)} = \frac{T_{seq}(N,1) + P\cdot T_{par}(N,1)}{T(N,1)} = \alpha + (1-\alpha)P$$

$$= P + (1-P)\alpha$$

Da cui

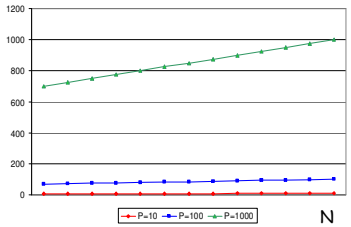
$$\lim_{N\to\infty} S_P = P \qquad \lim_{N\to\infty} E_P = \frac{S_P}{P} = 1$$

Legge di Gustafson
(reinterpretazione della legge di Amdhal)

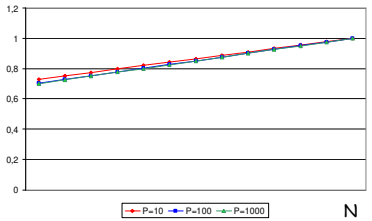
16

legge di Gustafson

La legge di Gustafson afferma che se la dimensione del problema aumenta sufficientemente, qualunque efficienza puo' essere ottenuta con un numero qualunque di processori



Speed Up



Efficienza

17