Prova Intermedia del 3 Novembre 2015 – traccia A

Domanda 1. Spiegate in poche righe perché non è più possibile aumentare la velocità dei processori aumentando semplicemente la velocità di clock. Illustrate inoltre la legge di Moore.

Domanda 2. Rispondete alle seguenti domande, motivando la risposta.

- Un programma contiene una funzione m che non può essere parallelizzata in cui viene speso il 40% del tempo di esecuzione. Qual è il limite allo speedup assumendo di avere un numero di processori arbitrariamente grande?
- Un programma contiene una funzione m che può essere parallelizzata ottenendo uno speedup 3 sulla sua stessa esecuzione. Quale frazione del tempo di esecuzione deve essere speso in m per dimezzare il tempo di esecuzione dell'intero programma?
- Per risolvere uno stesso problema due programmatori hanno sviluppato due diversi programmi: il programma $\mathcal{A}_1$ esegue work t_1 e ha una frazione di codice sequenziale $s_1 = 70\%$, il programma $\mathcal{A}_2$ esegue work t_2 e ha una frazione di codice sequenziale $s_2 = 30\%$. Supponete che $t_2 = 2\,t_1$:
  - se $\mathcal{A}_1$ e $\mathcal{A}_2$ sono eseguiti su una stessa piattaforma con P processori, quale esecuzione sarà più veloce?
  - se $\mathcal{A}_1$ è eseguito su una piattaforma con P processori e $\mathcal{A}_2$ su una piattaforma con $\frac{P}{2}$ processori, quale esecuzione sarà più veloce?

  Se la risposta dipende dal numero dei processori, spiegate sotto quali ipotesi l'esecuzione del programma $\mathcal{A}_1$ è più / meno veloce dell'esecuzione del programma $\mathcal{A}_2$.

Domanda 3. Considerate il codice per la somma degli elementi di un array visto a lezione, modificando il metodo `run` come segue:

```
public void run() {
    for (int i=lo; i<hi; i++)
        ans += processValue(arr[i]);
}
```

Anziché aggiungere direttamente al campo `ans` il valore `arr[i]`, il metodo `run` somma un valore calcolato a partire da `arr[i]` tramite il metodo `processValue`. Il metodo `processValue` ha però tempi di esecuzione diversi nelle varie chiamate. In particolare in questo esercizio:

- sia t il tempo di esecuzione della chiamata `processValue(arr[0])`;
- per ogni $i \geq 0$, la chiamata `processValue(arr[i])` richiede tempo $(i+1) \cdot t$.

Immaginate ora di suddividere la computazione tra due soli thread, come mostrato di seguito:

```
static int sum(int[] arr) throws java.lang.InterruptedException {
  int M = ... ;
  SumThread leftThread  = new SumThread(arr,0,M);
  SumThread rightThread = new SumThread(arr,M,arr.length);
```

```
    leftThread.start();
    rightThread.start();

    leftThread.join();
    rightThread.join();

    return leftThread.ans+rightThread.ans;
  }
```

Rispondete alle seguenti domande:

- Cosa comporta scegliere `M=arr.length/2`, come abbiamo fatto nel caso della somma? I tempi di esecuzione dei due thread sono bilanciati?
- Come sceglieresteste `M` per bilanciare meglio il carico? Date un'intuizione e calcolate una stima del valore ottimale per `M`.

  Suggerimento. Assumete che il tempo di esecuzione di un thread sia dato dalla somma dei tempi di esecuzione delle chiamate al metodo `processValue` effettuate dal thread stesso e ricordate la somma di Gauss:

  $$\sum_{x=1}^{n} x = \frac{n(n+1)}{2} \approx n^2$$

  Calcolate il tempo di esecuzione di ciascuno dei due thread ed imponete che i due tempi si eguaglino.
- Se avessimo usato l'approccio basato su *divide et impera* (senza cutoff sequenziale), in questo esempio sarebbe stato altrettanto cruciale bilanciare bene il carico? Ovvero, avremmo potuto scegliere `M=arr.length/2`? Motivare la risposta.

Domanda 4. Dopo aver introdotto il problema delle somme prefisse, descrivete l'algoritmo di Hillis & Steele ed analizzatene work e span.