

Nome e Cognome \_\_\_\_\_ Matricola \_\_\_\_\_

## Impianti di Elaborazione LS – Seconda Prova in Itinere

28/06/2004

**Nota preliminare:** specificare e giustificare le ipotesi utilizzate per la risoluzione degli esercizi e utilizzare una buona approssimazione nei conti (**con almeno 4 cifre decimali**) per evitare il propagarsi di errori di arrotondamento

### **SVOLGERE UN ESERCIZIO PER FOGLIO**

1) Considerare il sistema utilizzato da una società di assicurazioni per offrire servizi Web. Il sistema è costituito da un server Web, che ha la funzione di front-end verso l'esterno, un application server, che riceve le richieste dal server Web, e un database server, che gestisce il database della società. Ai fini del modello, il server Web è rappresentato da un centro di servizio W, l'application server da un centro di servizio AP, mentre il database server è rappresentato da 4 centri di servizio: un processore P e tre dischi D1, D2 e D3. Le richieste che arrivano al sistema si suddividono in due classi. Le richieste di classe 1 arrivano dall'esterno al server Web W, da cui si dirigono all'application server AP e quindi al database server, dove iniziano la loro elaborazione al processore P. In uscita dal processore P, le richieste accedono ad uno dei tre dischi oppure tornano all'application server. Dopo ogni accesso a disco le richieste tornano al processore P. Il numero di accessi al disco D1 è pari a 3, il numero di accessi al disco D2 è pari a 2 e il numero di accessi al disco D3 è pari a 1. Le richieste escono dal database server dopo una visita al processore P. In uscita dal database server le richieste si dirigono all'application server AP e poi al server Web W, dove terminano la loro elaborazione ed escono dal sistema.

Le richieste di classe 2 arrivano dall'esterno e accedono al server Web W, da cui si dirigono al database server, iniziando la loro elaborazione con una visita al processore P. In uscita dal processore P, le richieste accedono al disco D2, da cui tornano al processore P, procedono quindi verso il server Web W, dove terminano la loro elaborazione ed escono dal sistema. Supporre che le richieste di classe 1 e di classe 2 arrivino al sistema con frequenza  $\lambda_1 = 3.6$  richieste/minuto e  $\lambda_2 = 0.6$  richieste/minuto, rispettivamente. Supporre inoltre che:

$S_{1W} = 3$  sec,  $S_{1AP} = 4$  sec,  $S_{1P} = 1.4$  sec,  $S_{1D1} = 2.8$  sec,  $S_{1D2} = 4.8$  sec,  $S_{1D3} = 6.8$  sec

$S_{2W} = 6$  sec,  $S_{2P} = 1.8$  sec,  $S_{2D2} = 8.5$  sec.

- a) disegnare per ciascuna delle due classi la topologia del modello, specificando le probabilità associate a ciascun ramo
- b) determinare il bottleneck di ciascuna classe e il bottleneck del sistema
- c) calcolare
- c1) throughput totale del sistema

- c2) tempo di risposta delle richieste di classe 2
  - c3) numero di richieste di classe 1 in coda al processore P
  - c4) tempo totale che le richieste di classe 1 spendono in coda al centro di servizio bottleneck del sistema
  - d) calcolare, esprimendolo tramite la legge di Little, il numero di richieste di classe 2 nel sistema
- Supporre di inserire una nuova classe (chiusa) di richieste rappresentata da un programma che verifica la consistenza dei dati presenti sul database server; il programma inizia la sua elaborazione dal processore P, quindi fa un accesso a ciascuno dei dischi; dopo ogni accesso ai dischi torna al processore. Il programma termina con una visita al processore. Supporre:  $S_{3P} = 20$  sec.,  $S_{3D1} = 3$  sec.,  $S_{3D2} = 35$  sec.,  $S_{3D3} = 25$  sec e  $N=1$ .
- e1) disegnare la topologia della nuova classe di richieste
  - e2) determinare il bottleneck della nuova classe
  - e3) si riesce a stabilire il bottleneck del sistema senza risolvere il modello? Perché?
  - e4) risolvere il modello e calcolare l'utilizzo per classe di ciascuno dei centri di servizio
  - e5) calcolare il peggioramento (miglioramento?) del tempo di risposta delle richieste di classe 2
  - e6) calcolare la variazione del throughput del sistema dovuto alla presenza della nuova classe

**FACOLTATIVO** (da svolgere SOLO a compito completato): a partire dal modello aperto iniziale, distribuire gli accessi ai tre dischi da parte delle richieste di classe 1 in modo da ottenere un carico bilanciato sui dischi; calcolare le variazioni del tempo che le richieste di classe 1 spendono nel database server e del numero di richieste di classe 1 e di classe 2 presenti nel database server

- 2) Considerare il modello di un sistema in cui sono presenti i terminali utente, un processore P e due dischi D1 e D2. Gli utenti sono caratterizzati da un tempo di think pari a 9 secondi. Le richieste inviate dai terminali utente accedono quattro volte il processore P, due volte il disco D1 e una volta il disco D2, prima di tornare ai terminali utente. I tempi di servizio sono:  $S_P = 0.2$  sec,  $S_{D1} = 3.2$  sec e  $S_{D2} = 4$ . sec.
- a) applicare le tecniche di analisi asintotica e calcolare e disegnare gli asintoti di throughput e tempo di risposta al variare del numero N di utenti
  - b) calcolare il valore  $N^*$  punto di intersezione tra gli asintoti
  - c) verificare e discutere se è ammissibile avere un throughput di 0.15 richieste/sec quando nel sistema sono presenti  $N=10$  utenti
  - d) determinare gli asintoti bilanciati di throughput e tempo di risposta