

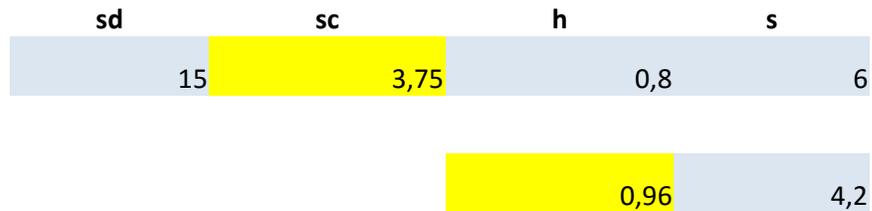
# Dischi e CPU

Alcuni esercizi sulle prestazioni

# Esercizio 1

- Un disco possiede un tempo medio di servizio pari a 15 ms. nel caso in cui si escluda la cache;  
se invece la cache è attiva, con una hit ratio (probabilità di trovare il dato in essa) pari a 80%, il tempo di servizio è di 6 ms.
- Calcolare il valore che deve avere tale probabilità perché il tempo di servizio diventi 4.2 ms. a parità delle altre condizioni.

# Esercizio 1 soluzione



$$s = sd \times (1-h) + sc \times h$$

$$sc = (6 - 15 \times 0.2) / 0.8 = 3.75$$

$$h = (15 - 4.2) / (15 - 3.75) = 0.96$$

$s$  = tempo di servizio

$sd$  = tempo di servizio disco

$sc$  = tempo di servizio cache

$h$  = probabilità di hit

## Esercizio 2

- Un dispositivo hard disk ha le seguenti caratteristiche di funzionamento:
- rotazione: 7200 rpm; (rotazioni per minuto)
- tempo medio seek: 7.2 ms.;
- tempo medio di trasferimento di un blocco di dati da disco: 0.2 ms.
- tempo medio di trasferimento da cache: 0.05 ms.
- 1. sapendo che il “seek” ha luogo solo in 1/3 delle operazioni, **calcolare: a) *il tempo medio di servizio del disco nel caso in cui la cache sia disabilitata.***
- 2. se si introduce la cache, **calcolare: b) la “miss rate”** che porterebbe a una riduzione del 60% del tempo di servizio.

# Esercizio 2

## soluzione

rpm	seek tm	transfer	cache	seek prob	miss prob	serv
7200	7,2	0,2	0,05	0,333333	1	6,766667
					0,395533	2,706667

rot delay  
4,166667

riduz. serv  
0,6

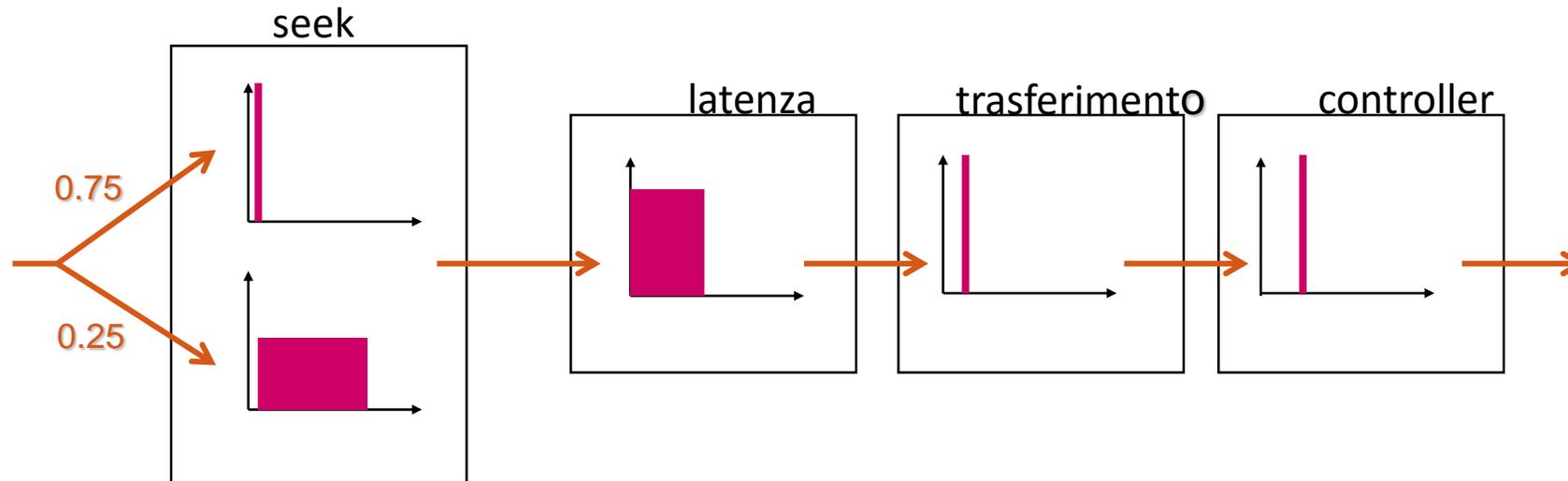
serv1

serv2

miss prob

$$\begin{aligned} \text{rot\_delay} &= 60 \times 1000 / (\text{rpm} \times 2) \\ \text{serv1} &= \text{rot\_delay} + \text{seek\_tm} \times \text{seek\_prob} + \text{transfer} \\ \text{serv2} &= \text{serv1} \times (1 - \text{miss}) + \text{cache} \times \text{miss} \\ \text{miss\_prob} &= (\text{serv2} - \text{cache}) / (\text{serv1} - \text{cache}) \end{aligned}$$

# Riassunto grafico di un caso analogo



$t_1 = 0 ; V_1 = 0$   
 $t_2 = 6 ; V_2 = 10.45$

$T(\text{seek}) = 1.5$   
 $V(\text{seek}) = 9.36$   
 $C = 2.04$

$T(\text{latenza}) = 3$   
 $V(\text{latenza}) = 3$   
 $C = 0.58$

$T(\text{trasf.}) = 0.01$   
 $V(\text{trasf.}) = 0$   
 $C = 0$

$T(\text{contr.}) = 0.2$   
 $V(\text{contr.}) = 0$   
 $C = 0$

il fenomeno complessivo ha un tempo medio  $s = \sum T = 4.71$   
e varianza  $= \sum V = 12.36 ; C = 0.747$

# Digressione sulla media

Il tempo di risposta  $R=g(X)$  in funzione del traffico  $X$  ha la concavità rivolta verso l'alto.

Il punto  $(X,R)$ , media delle osservazioni, non si trova sulla curva  $g(X)$ .

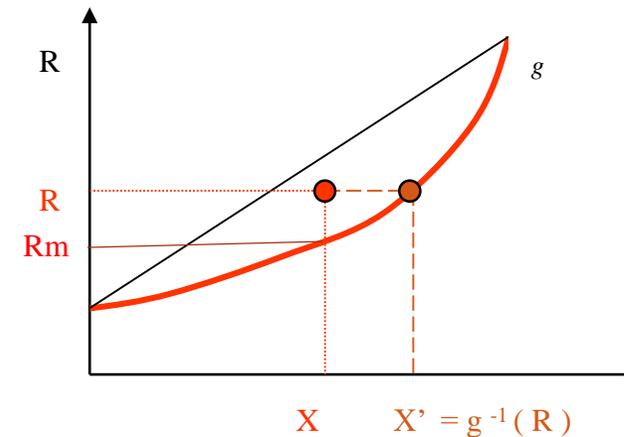
$X'$ : flusso di transazioni che darebbe luogo al tempo medio di risposta misurato  $R$

( $X = X'$  solo nel caso di  $g$  funzione lineare).

Durante il periodo di osservazione il punto rappresentativo del fenomeno si muove sulla curva  $R = g(X)$  al variare di  $X$

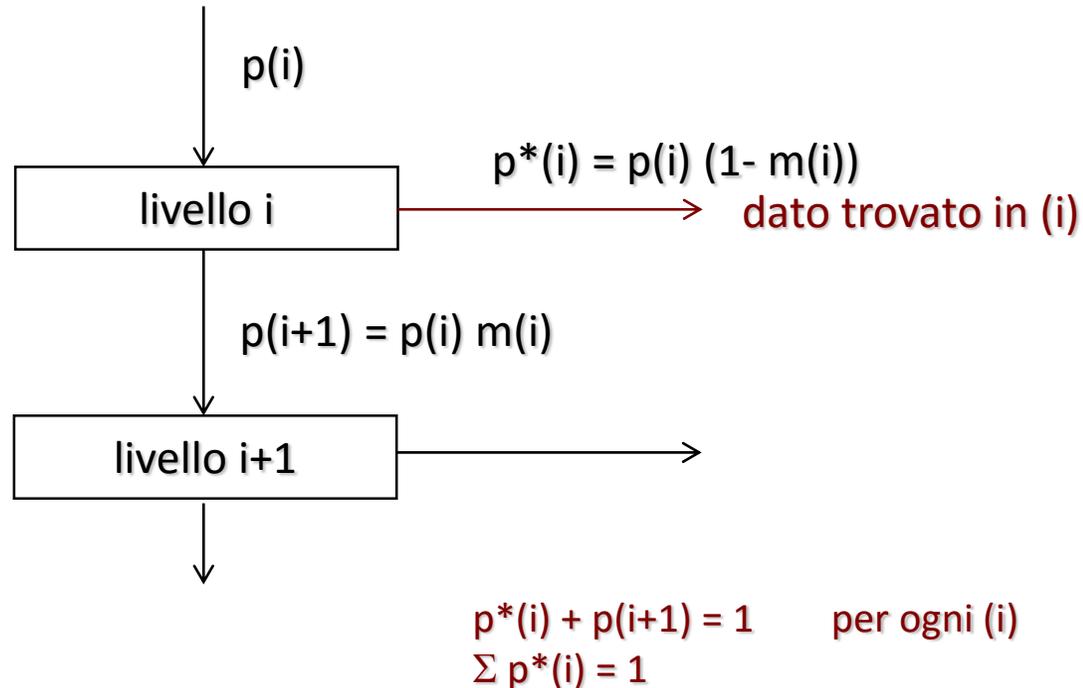
Se tutte le misure corrispondessero al tasso medio  $X$ , il tempo medio  $R_m$  apparirebbe al grafico di  $g$  e sarebbe  $R_m < R$

→ Ai fini del tempo medio di risposta conviene che il traffico si distribuisca uniformemente fra i componenti



$X'$  è il tasso di transazioni che darebbe luogo al tempo di risposta ottenuto come media delle misure

# Calcolo del tempo medio in una gerarchia di memoria



$m(i)$ : probabilità di «miss» al livello (i)

$p(i)$ : probabilità di ricerca nel livello (i) essendo al livello (i-1)

$p^*(i)$ : probabilità di fermarsi nella ricerca al livello (i)

# Esercizio 3

Una gerarchia di memoria è composta da sette livelli, si calcoli il tempo medio di accesso al dato, considerando i seguenti dati:

livello	Tempo di accesso	miss rate	
1	1	1E-1	
2	1	5E-2	
3	10	2E-2	
4	1E2	1E-1	
5	1E7	2E-2	
6	5E7	2E-2	
7	4E8	0	

# Esercizio 3

# soluzione

livello			prob1	p1 x t	acc		
	acc time	miss rate			prob2	time2	p2 x t2
1	1,00E+00	1,00E-01	1,00E+00	1,00E+00	9,00E-01	1,00E+00	9,00E-01
2	1,00E+00	5,00E-02	1,00E-01	1,00E-01	9,50E-02	2,00E+00	1,90E-01
3	1,00E+01	2,00E-02	5,00E-03	5,00E-02	4,90E-03	1,20E+01	5,88E-02
4	1,00E+02	1,00E-01	1,00E-04	1,00E-02	9,00E-05	1,12E+02	1,01E-02
5	1,00E+07	2,00E-02	1,00E-05	1,00E+02	9,80E-06	1,00E+07	9,80E+01
6	5,00E+07	2,00E-02	2,00E-07	1,00E+01	1,96E-07	6,00E+07	1,18E+01
7	4,00E+08	0,00E+00	4,00E-09	1,60E+00	4,00E-09	4,60E+08	1,84E+00
				112,76			112,76

acc time(i): tempo di accesso relativo al livello (i)

time2(i): tempo totale di accesso fino al livello (i) compreso

# «Potenza» di una CPU

- Definiamo come **potenza della CPU\_A**, relativamente ad un processore preso come unità base (CPU\_U), il rapporto inverso fra i tempi di esecuzione in CPU di un particolare carico:
  - Questa definizione ha senso se le altre risorse del sistema (dischi etc.) non costituiscono un «collo di bottiglia»
- **$P(A/U) = T_U/T_A$** 
  - Tale rapporto dipende sia dalle CPU sia dal tipo di carico cioè:  
 $P_1(A/U) = T_{1U}/T_{1A}$ : potenza della CPU A rispetto al tipo di carico 1
  - Perciò se avessimo tre tipi di carico con tempi, misurati su U, di ripartizione:  
 $(\alpha_1, \alpha_2, \alpha_3); \quad \alpha_1 + \alpha_2 + \alpha_3 = 1$
  - $P(A/U)$  dipende oltre che dal tipo di carico dai valori  $\alpha$

# Esercizio 4: «potenza» di una CPU

$$\bullet P(A/U) = \frac{\alpha_1 T u_1 + \alpha_2 T u_2 + \alpha_3 T u_3}{\frac{\alpha_1 T u_1}{P_1} + \frac{\alpha_2 T u_2}{P_2} + \frac{\alpha_3 T u_3}{P_3}}$$

- $\Rightarrow$  la potenza media è la media armonica delle potenze con pesi uguali alla ripartizione dei carichi  $\times$  i tempi (misurati sulla CPU base):  $\alpha_i \times T U_i$

CPU	U	A			
carichi	tempi	potenza	rip. carichi	pesi	
1	0,06296	0,04835	1,302172	5	0,001112
2	2,713	2,091	1,297465	3	0,028742
3	137,36	102,16	1,344558	2	0,970146

1,343108

# «Potenza» di un sistema

- Il numero di esecuzioni nell'unità di tempo di una applicazione significativa per il sistema può essere usata come indice della sua potenza
- Esempio pratico:
  - definiamo come *potenza* il *numero di blocchi elaborati al secondo da un programma* che ripete indefinitamente il seguente ciclo di tre passi:
    - **lettura** di un blocco di **8KB** su disco
    - **elaborazione** in CPU (30 M cicli)
    - **scrittura** di un altro blocco di 8KB su disco
  - il programma è l'unico processo attivo
  - non c'è alcuna sovrapposizione fra operazioni su disco e CPU

# Esercizio 5: «Potenza» di un sistema

- caratteristiche **CPU**
  - frequenza 5 GHz
  - tempo di elaborazione in CPU per ciclo di programma  
 $= 30 \times 10^6 / 5 \times 10^9 = 0.006 \text{ s.}$
- caratteristiche operazione **I/O disco**

	caratteristiche	tempo (ms)	tempo (ms)		
			disco	cache	
seek		8	8	0	
RPM	15000	2	2	0	
trasf MB/sec	80	0,0977	0,0977	0	
cache MB/sec	320	0,0244	0	0,0244	
controller		1	1	1	
serv. time (ms)			11,0977	1,0244	<b>7,0684</b>

latenza =  $(60000/15000)/2 = 2 \text{ ms.}$

trasf (disco) =  $8 \times 1000 / (80 \times 1024)$   
 $= 0.0977 \text{ ms.}$

trasf (cache) =  $8 \times 1000 / (320 \times 1024) = 0.0244 \text{ ms.}$

tempo di servizio I/O =  $0.6 \times 11.0977 + 0.4 \times 1.0244$   
 $= 7.0684 \text{ ms.}$

(si assume che il dato si trovi nella cache nell' 80% delle letture,  
 pari al 40% del totale delle operazioni)

# Esercizio 5

- l'esecuzione di un ciclo di elaborazione richiede perciò:
- $6 + 2 \times 7.0684 = 20.1367$  ms.
- la potenza è quindi:
- $1000 / 20.1367 = 49.66$  blocchi/sec.
- Utilizzo CPU =  $6 / 20.1367 = 29.80\%$
- Utilizzo disco =  $2 \times 7.0684 / 20.1367 = 100 - 29.80 = 70.20\%$