

# Performance Modeling of Computer Systems and Networks

*Prof. Vittoria de Nitto Personè*

## Abstract Priority

Università degli studi di Roma Tor Vergata  
Department of Civil Engineering and Computer Science Engineering

Copyright © Vittoria de Nitto Personè, 2021  
<https://creativecommons.org/licenses/by-nc-nd/4.0/>



1

Analytical models  
*abstract priority*

## QoS management

- Service provider
- Traffic flows with different QoS
- QoS: mean response time

prof. Vittoria de Nitto Personè

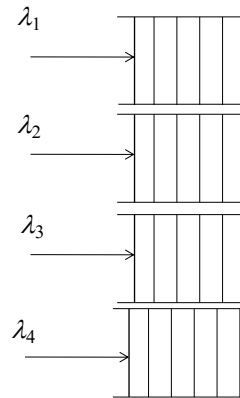
2

2

Abbiamo tempo medio uguali per tutti. (Abstract)

Analytical models  
abstract priority

## M/M/1 – NP\_priority



$$\lambda = \sum_{i=1}^r \lambda_i \quad \text{uniform partition: } \lambda_i = \frac{\lambda}{4}, p_i = \frac{1}{4}$$

stessa occupazione per ogni classe

$$\rho = 0.1, 0.2, 0.4, 0.6$$

Prob di essere di  
classe "k" = 1/4

$$E(T_{Q_i})^{NP\_priority} = \frac{\frac{\lambda}{2} E(S^2)}{\left(1 - \sum_{i=1}^k \rho_i\right) \left(1 - \sum_{i=1}^{k-1} \rho_i\right)}$$

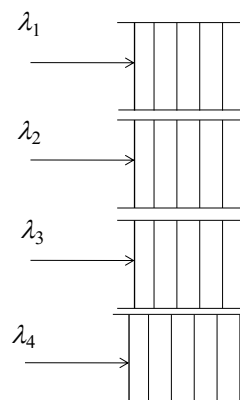
prof. Vittoria de Nitto Personè

3

3

Analytical models  
abstract priority

## NP priority



$$E(T_{Q_1}) = \frac{\rho E(S)}{(1 - \rho_1)}$$

$$E(T_{Q_2}) = \frac{\rho E(S)}{(1 - (\rho_1 + \rho_2))(1 - \rho_1)}$$

$$E(T_{Q_3}) = \frac{\rho E(S)}{(1 - (\rho_1 + \rho_2 + \rho_3))(1 - (\rho_1 + \rho_2))}$$

$$E(T_{Q_4}) = \frac{\rho E(S)}{(1 - \rho)(1 - (\rho_1 + \rho_2 + \rho_3))}$$

prima classe va  
sempre meglio, ovvio.

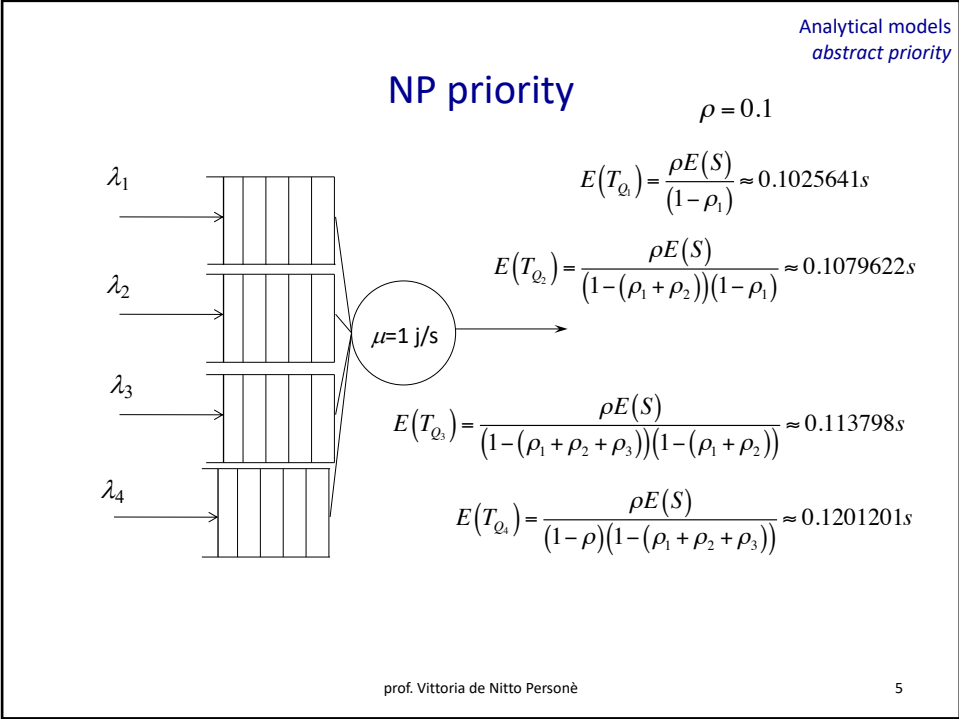
l'ultima va sicuramente  
peggio.

prof. Vittoria de Nitto Personè

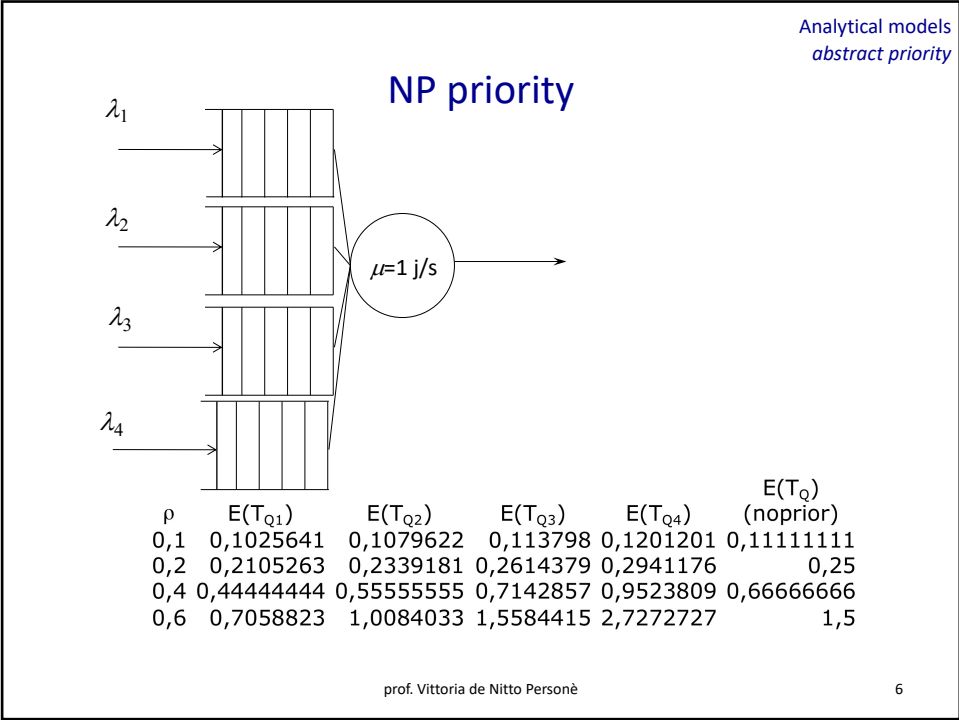
4

4

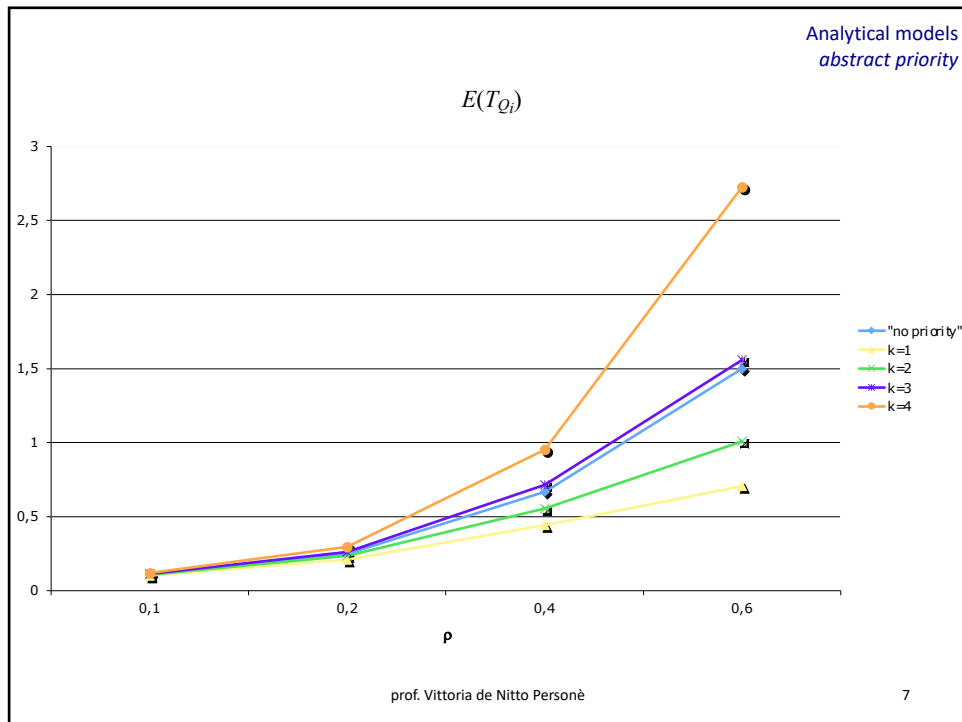
Il numeratore vede sempre l'occupazione di tutte le classi.



5



6



divide in due gruppi i vari grafici di "k".

7

Analytical models  
priority scheduling

Goals:

- given a QoS requirement, decide if adopt priority classes
- note that if the policy is non-size-based, we can reason just in terms of waiting time

Mean service demand: 0.4 s

QoS requirement  
the waiting time should not exceed the service demand, in particular:  
the service provider will not incur in penalties if  $T_Q \leq 0.45$  s;  
the service provider will gain revenue if  $T_Q < 0.4$  s

By simple "costless" analysis we can offer good insights

prof. Vittoria de Nitto Personè

8

8

$$E(S) = 0.4 \text{ s}$$

Low load medium load high load

$$\rho = 0.4 \quad 0.6 \quad 0.8$$

$$\lambda = 1 \quad 1.5 \quad 2 \text{ job/s}$$

$$E(T_Q) = 0.26 \quad 0.6 \quad 1.6 \text{ job/s} \quad \text{without priority classes}$$

manca grafico:

in "medium load", con caso 30% e 70%, le prestazioni vanno meglio per entrambi. Basta alternare le percentuali, e le prestazioni vanno entrambe peggio. Perché 30% e 70% meglio?

Alla classe 1 do un privilegio maggiore, ma a "pochi". La seconda classe è superata dal 30% di quelli di classe 1, quindi superata da "pochi".

Nel caso 70% e 30% va peggio perché do la priorità a troppi job.

high load  $\rho = 0.8$

not penalties if  $T_Q \leq 0.45 \text{ s}$

gain revenue if  $T_Q < 0.4 \text{ s}$

$$p_1 = 0.36, p_2 = 0.64$$

$$p_1 = 0.22, p_2 = 0.78$$

$$E(T_{Q1}) = 0.4494 \text{ s}$$

$$E(T_{Q1}) = 0.3883 \text{ s}$$

$$E(T_{Q2}) = 2.2472 \text{ s}$$

$$E(T_{Q2}) = 1.9417 \text{ s}$$

Nel secondo caso, sia  $E[T_{q1}]$  che  $E[T_{q2}]$  sono migliori del primo caso. Tuttavia, nel secondo caso ho ricompensa solo nel 22% dei casi, nel primo non ricevo penalità nel 36%. Nel secondo caso ho tempi medi migliori, ma con percentuali minori.

metto requisito di qualità su prima classe, vedo quale valore di  $p_1$  ci da tempo di attesa minore di 0.4  
 $\rho_{01} = p_1 \cdot \rho$ ,  $\rho \cdot E[s] / (1 - p_1 \cdot \rho) \leq 0.45$ , trovo  $p_1$  che soddisfa.  $0.45 = QoS$

$$E(T_Q)_{glob} = E(T_Q)_{KP} = 1.6 \text{ s}$$

## Euristica per la ripartizione in classi di priorità astratta

coda singola  
 $\rho = 0.92, E(S) = 1 \text{ j/s} \quad E(T_Q) = 11.5 \text{ s}, E(T_S) = 12.5 \text{ s}$

60%, 25 %, 15 %,	15 %, 25 %, 60%,
$E(T_{Q1}) = 2.05357 \text{ s}$	$E(T_{Q1}) = 1.067285 \text{ s}$
$E(T_{Q2}) = 9.42005 \text{ s}$	$E(T_{Q2}) = 1.688743 \text{ s}$
$E(T_{Q3}) = 52.75229 \text{ s}$	$E(T_{Q3}) = 18.196203 \text{ s}$
$E(T_{S1}) = 3.05357 \text{ s}$	$E(T_{S1}) = 2.067285 \text{ s}$
$E(T_{S2}) = 10.42005 \text{ s}$	$E(T_{S2}) = 2.688743 \text{ s}$
$E(T_{S3}) = 53.75229 \text{ s}$	$E(T_{S3}) = 19.196203 \text{ s}$

A sinistra, migliore classe 1 e classe 2 (migliore nell'85%) rispetto coda unica.  
 Il problema è che Q3 cresce molto.

prof. Vittoria de Nitto Personè

11

11

arrivi random/ exp (quite variable) / 1 servente

Analytical models  
abstract priority

M/M/1

Arrival flow: random, rate 0.8 job/s

Service process: quite variable rate 1 job/s

QoS requirements

the response time should not exceed twice the

service demand, in particular:

the service provider will not incur in

penalties if  $T_S \leq 4 \text{ s}$ ;

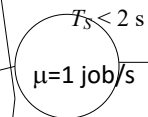
the service provider will gain revenue if

$\lambda = 0.8 \text{ job/s}$

20%

30%

50%



$\rho = 0.8$

con coda singola  $E(T_Q) = \frac{\rho E(S)}{1 - \rho} = \frac{0.8}{0.2} = 4 \quad E(T_S) = 5 \quad \text{non rispetto i requisiti.}$

prof. Vittoria de Nitto Personè

12

12

M/M/1

Not penalties if  $T_S \leq 4$  s;  
gain revenue if  $T_S < 2$  s

	NP	P (prelazione)	NP	P
class	$E(T_Q)$		$E(T_S)$	
1 - 20%	0.9523809523809524	0.19047619047619052	1.9523809523809526	1.1904761904761905
2 - 30%	1.5873015873015874	0.7936507936507937	2.5873015873015874	1.9841269841269842
3 - 50%	6.666666666666669	6.666666666666669	7.666666666666669	8.333333333333336
global	4.000000000000001	3.6095238095238105	5.000000000000001	5.000000000000001

La classe 3, ha tempo  $E[T_Q]$  NP e P uguali, essendo ultima classe non fa prelazione.  
La classe 1 e classe 2, facendo prelazione, hanno guadagno rispetto alla variante senza prelazione.  
Globalmente ho guadagno rispetto a KP, in virtù della prelazione. Con KP ho tempo globale 4.  
Passiamo ai tempi di risposta: perdo il guadagno perchè devo considerare il tempo di servizio che aumenta per i job che vengono buttati fuori (colpa della MEMORYLESS).  
Vedendo  $E[T_S]$  NP, in due casi riesco a stare sotto 4s, con prelazione ho anche guadagno.

prof. Vittoria de Nitto Personè

13

13

M/M/1

Not penalties if  $T_S \leq 4$  s;  
gain revenue if  $T_S < 2$  s

	NP	P	NP	P
class	$E(T_Q)$		$E(T_S)$	
1 - 20%	0.9523809523809524	0.19047619047619052	1.9523809523809526	1.1904761904761905
2 - 30%	1.5873015873015874	0.7936507936507937	2.5873015873015874	1.9841269841269842
3 - 50%	6.666666666666669	6.666666666666669	7.666666666666669	8.333333333333336
global	4.000000000000001	3.6095238095238105	5.000000000000001	5.000000000000001

prof. Vittoria de Nitto Personè

14

14

Analytical models  
abstract priority

M/M/1

Not penalties if  $T_S \leq 4$  s;  
gain revenue if  $T_S < 2$  s

	NP	P	NP	P
class	$E(T_Q)$		$E(T_S)$	
1 - 20%	0.9523809523809524	0.19047619047619052	1.9523809523809526	1.1904761904761905
2 - 30%	1.5873015873015874	0.7936507936507937	2.5873015873015874	1.9841269841269842
3 - 50%	6.666666666666669	6.666666666666669	7.666666666666669	8.333333333333336
global	4.000000000000001	3.6095238095238105	5.000000000000001	5.000000000000001

prof. Vittoria de Nitto Personè15

15

Aggiungiamo una variabilità maggiore

Analytical models  
abstract priority

No priority classes  
 $g(0.1)=4.5556$ ,  
 $E(T_Q)=11.1111$  s,  $E(T_S)=12.1111$  s

kp unica coda

M/H<sub>2</sub>/1 iperesponenziale

Not penalties if  $T_S \leq 4$  s;  
gain revenue if  $T_S < 2$  s

	NP	P	NP	P
class	$E(T_Q)$		$E(T_S)$	
1 - 20%	2.645502645502645	0.5291005291005291	3.645502645502645	1.529100529100529
2 - 30%	4.409171075837742	2.204585537918871	5.409171075837742	3.3950617283950617
3 - 50%	18.51851851851852	18.51851851851852	19.51851851851852	20.185185185185187
global	11.111111111111111	10.026455026455027	12.111111111111111	11.416931216931218

piccolo guadagno con la prelazione, nella memoryless non c'era.

p=0.1 → 10% E(S)=5 s  
90% E(S)=0.55555 s

prof. Vittoria de Nitto Personè16

16

E[S] circa 1, coincide con quella di partenza. Stessa media globale del servizio, sennò non ha senso con un carico diverso.

solo con P, E[TS] sta sotto i 4s, ma non riesco a stare con entrambi le classi sotto 2s per avere guadagno.



## Altro approccio per valutare il REVENUE

Mean service demand (expo): 0.4 s

QoS requirement

the waiting time (average) should not exceed 0.1 s, in particular:

the service provider will gain  $c_1$  for each service within QoS

the service provider will pay  $c_2$  for each service violates QoS

quantità soddisfano\*guadagno - quantità non soddisfano\*perdita

$$R = p_1 c_1 - p_2 c_2$$

Abstract-P  $\rightarrow$  max R

prof. Vittoria de Nitto Personè

17

17

esempio:

$$E(S) = 0.4 \text{ s}, \lambda = 0.8 \text{ j/s}, \rho = 0.32$$

trovato con la formula che abbiamo visto qualche slide fa.

$$p_1 = 0.6, p_2 = 0.4, c_1 = 5, c_2 = 3 \rightarrow R = 2.2$$

costanti

$$E(T_{Q1}) = 0.095 \text{ s}, E(T_{S1}) = 0.495 \text{ s}$$

$$E(T_{Q2}) = 0.233 \text{ s}, E(T_{S2}) = 0.728 \text{ s}$$

prof. Vittoria de Nitto Personè

18

18