

# Sistemi e Applicazioni Cloud

Appello del 20 febbraio 2025 [Tempo consegna: 2h 30m]

## Parte 1: rete base

Si usi un simulatore per studiare il comportamento di un sistema in grado di parallelizzare il traffico su diversi nodi.

Il sistema è mostrato nella figura.

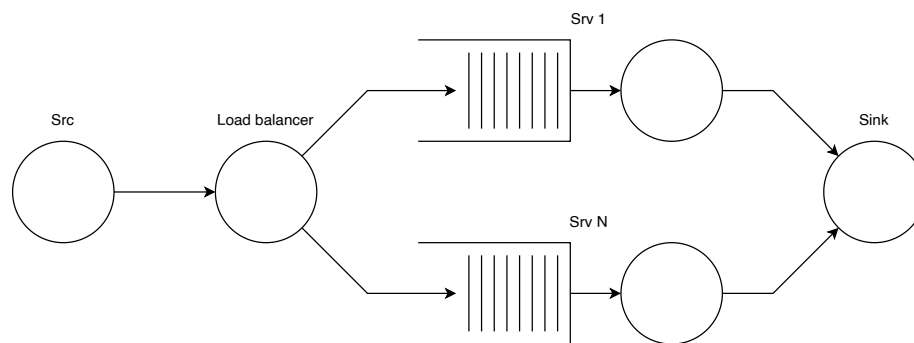


Figure 1: Modello di rete

Il carico in ingresso è  $\lambda = 100$  richieste al secondo e viene ripartito equamente tra gli  $N$  server (politica *round-robin* o *random* a piacere). I server hanno capacità di servizio  $\mu_1 = 10$  richieste/sec. Il tempo di servizio segue una distribuzione lognormal con coefficiente di variazione  $cv = 3$ . Il processo di servizio delle richieste è vincolato ad un SLA sul tempo di risposta medio  $T_r$  che deve restare al di sotto di 250 ms.

Testare il tempo di servizio per  $N = 20$  indicando anche l'intervallo di confidenza del 65% [ $\approx 600ms$ ].

$N$	$T_r$	$\pm$ CI
20		

## Parte 2: dimensionare il bilanciamento

Identificare mediante la teoria delle reti di code il valore di  $N$  tale per cui il requisito di SLA soddisfatto

$N$	$T_r$

### Parte 3: verifica

Eeguire un'analisi del tempo di risposta per un range di valori di  $N \in [15, 20, 25, 30, 35, 40, 45, 50]$ .

$N$	$T_r$	$\pm$ CI
15		
20		
25		
30		
35		
40		
45		
50		

Punto bonus: realizzare plot dei dati sulla base dell'esempio fornito

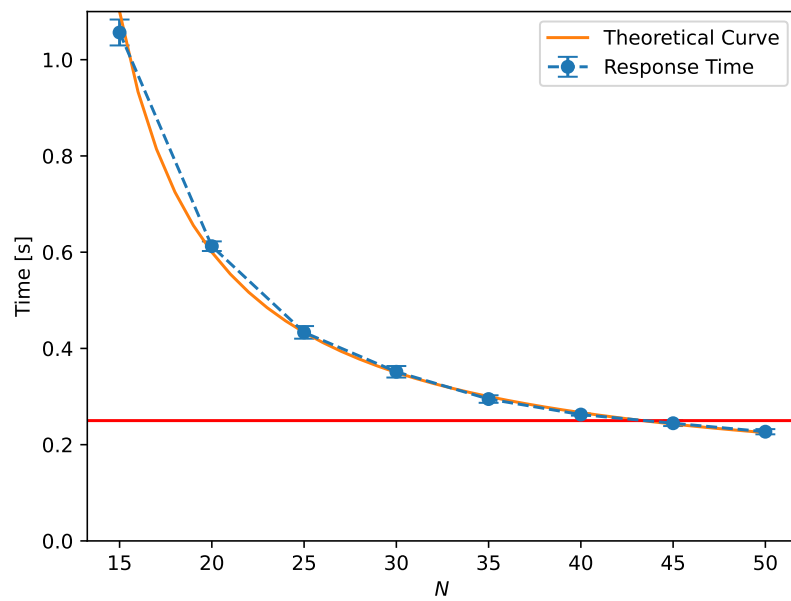


Figure 2: Plot