

Università degli Studi di Catania

Facoltà di Ingegneria

Corso di Laurea in Ingegneria Elettronica

Anno Accademico 2003/2004



Ottimizzazione di una rete con router RIO
mediante l'utilizzo di modelli fluidici

Modelli Fluidici

Uno dei problemi attuali per chi deve creare una rete di telecomunicazioni è il dimensionamento dei vari elementi della stessa. Con una metafora comune, ha un problema analogo per la scelta dei tubi e delle riserve idriche, chi deve creare una rete di acquedotti.

Nell'ottica del dimensionamento di reti TCP/IP, esistono due approcci differenti: l'approccio simulativo e quello analitico. Il primo è più semplice da sviluppare, ma ha l'inconveniente che per avere risultati affidabili necessita di molto tempo. Nel secondo, invece, il costo maggiore è dato dallo sviluppo di un modello di rete, basato sul comportamento medio, e che sia il più rispondente possibile alla rete da analizzare.

I modelli fluidici hanno il compito di modellizzare analiticamente una topologia prestabilita, supponendo che il flusso dei dati, normalmente "a pacchetti", sia invece continuo. Se immaginiamo il classico pacchetto "spalmato", nel suo tempo di vita, possiamo modellizzare la trasmissione su un link, proprio come l'invio in un tubo con una determinata capacità di un fluido ad una certa portata. In maniera analoga è possibile modellizzare l'accodamento o lo scarto di un certo numero di pacchetti. In Tale maniera, meccanismi del TCP come lo Slow Start o il Congestion Avoidance, possono essere rappresentati come eventi "continui", che possono verificarsi nell'insieme dei flussi contemporaneamente. È anche possibile studiare il comportamento di una rete che utilizza meccanismi di AQM come il RED o RIO.

Modello analitico

Il cuore di questa tipologia di simulazione è un set di equazioni di stato che descrivono in maniera continua i parametri della rete: Window Size, Round Trip Time, Drop Probability, Threshold Size, Stimed Queue Lenght e Real Queue Lenght.

L'equazione che descrive la dinamica delle code è la seguente:

$$\frac{dq_l(t)}{dt} = \begin{cases} -C_l + (1 - P_l^{(b)})\lambda_{NA}^{(q)}(t) & \text{per } q_l(t) > 0 \\ [-C_l + (1 - P_l^{(b)})\lambda_{NA}^{(q)}(t)]^+ & \text{per } q_l(t) \leq 0 \end{cases}$$

dove il primo termine rappresenta la velocità di decremento della coda, mentre la seconda parte rappresenta il rate di arrivo dei pacchetti non scartati e dove

$$[f(x)]^+ = \begin{cases} f(x) & \text{se } f(x) \geq 0 \\ 0 & \text{se } f(x) < 0 \end{cases}$$

Il Round Trip Time deve tener conto sia del tempo di propagazione dei vari link che del tempo di accodamento lungo le varie code nel path.

$$R_i(t) = \sum_{l \in L_i} \left(2d_l + \frac{q_l(t)}{C_l} \right)$$

Il modello però deve tener conto che il round trip time di riferimento è legato al tempo t (attuale) dalla seguente equazione implicita:

$$\tau_i + R_i(\tau_i) = t$$

che bisogna risolvere per tener conto che i pacchetti emessi τ_i secondi prima sono quelli che determinano la dinamica al tempo t .

La dinamica di ogni finestra di trasmissione si modella con un'equazione differenziale che sostanzialmente tiene conto dell'aumento della finestra (additive increase) e contemporaneamente della sua riduzione (multiplicative decrease):

$$\frac{dW_i(t)}{dt} \cong p_i^{(p)}(\tau_i) \frac{W_i(\tau_i)}{R_i(\tau_i)} \left[\left(1 - \frac{W_i(t)}{2}\right) Q(W_i(t)) - \frac{W_i(t)}{2} \right] + \begin{cases} \frac{(1 - p_i^{(p)}(\tau_i)) W_i(\tau_i)}{R_i(\tau_i)} & \text{per } W_i(t) < T_i(t) \\ \frac{(1 - p_i^{(p)}(\tau_i)) W_i(\tau_i)}{W_i(t) R_i(\tau_i)} & \text{per } W_i(t) \geq T_i(t) \end{cases}$$

Il primo termine dell'equazione tiene infatti conto del rate di dimezzamento della finestra ogni qualvolta arrivano tre ack duplicati o dell'abbattimento della stessa in caso di timeout. A tal proposito si considera l'espressione $Q(w)$ che modella la probabilità che, data la congestione window di dimensione w , dato per certo che è accaduta una perdita, questa sia dovuta allo scadere del Timeout. Nel seguito considereremo che una perdita dovuta a timeout accada se la transmission window ha un valore inferiore a tre o se il pacchetto perso di una sequenza di w è uno degli ultimi tre pacchetti. Matematicamente ciò si realizza ponendo:

$$Q(w) = \min\left(1, \frac{3}{w}\right)$$

Contemporaneamente il secondo termine invece modella le due fasi di crescita della window del TCP New Reno: lo Slow Start e il Congestion Avoidance. Nel primo caso infatti la finestra aumenta di W ogni RTT, dato per certo che non vi siano perdite; nel secondo caso la window aumenta di uno, ogni RTT, sempre dato per certo che non ci siano perdite.

Un altro parametro con una dinamica indipendente è il valore della soglia (Threshold) che separa la fase di Slow Start da quella di Congestion Avoidance; esso viene indicato con il termine $T_i(t)$ e l'equazione che ne regola l'andamento è la seguente:

$$\frac{dT_i(t)}{dt} \cong p_i^{(p)}(\tau_i) \frac{W_i(\tau_i)}{R_i(\tau_i)} \left[\frac{W_i(t)}{2} - T_i(t) \right]$$

Ogni volta che avviene una perdita la Threshold assume un valore pari alla metà della Congestion Window mentre, se non si hanno perdite, la variazione è nulla.

Per completare il modello bisogna aggiungere l'equazione che regola la dinamica della coda stimata dal meccanismo di AQM in oggetto e che, di fatto, determina la gestione della coda stessa. Tale equazione è:

$$\frac{dm_l(t)}{dt} = \begin{cases} \ln(1 - \alpha) \cdot [m_l(t) - q_l(t)] \cdot \lambda_l^{(Q)}(t) & \text{se } q_l(t) > 0 \\ \ln(1 - \alpha) \cdot m_l(t) \cdot C_l & \text{se } q_l(t) = 0 \end{cases}$$

Il RIO infatti utilizza tante maschere RED quanti sono i profili di QoS assegnati alle sorgenti. Queste maschere fanno corrispondere alla stima della coda una certa probabilità di scarto dei pacchetti sotto servizio. Il parametro α rappresenta il peso utilizzato dal filtro EWMA (Estimated Weighted Mobile Average) implementato nel meccanismo di controllo di congestione. Tale valore influenza la valutazione della coda stimata, che riesce a seguire l'andamento della coda reale tanto meglio quanto più α è vicino a uno a scapito di un minor grado di stabilità. Viceversa un valore di α

prossimo a zero ha la capacità di rendere maggiormente *smoothed* un flusso discontinuo, ma impedisce una tempestiva rivelazione della congestione se questa è improvvisa.

Topologia

La topologia oggetto del nostro studio è la seguente:

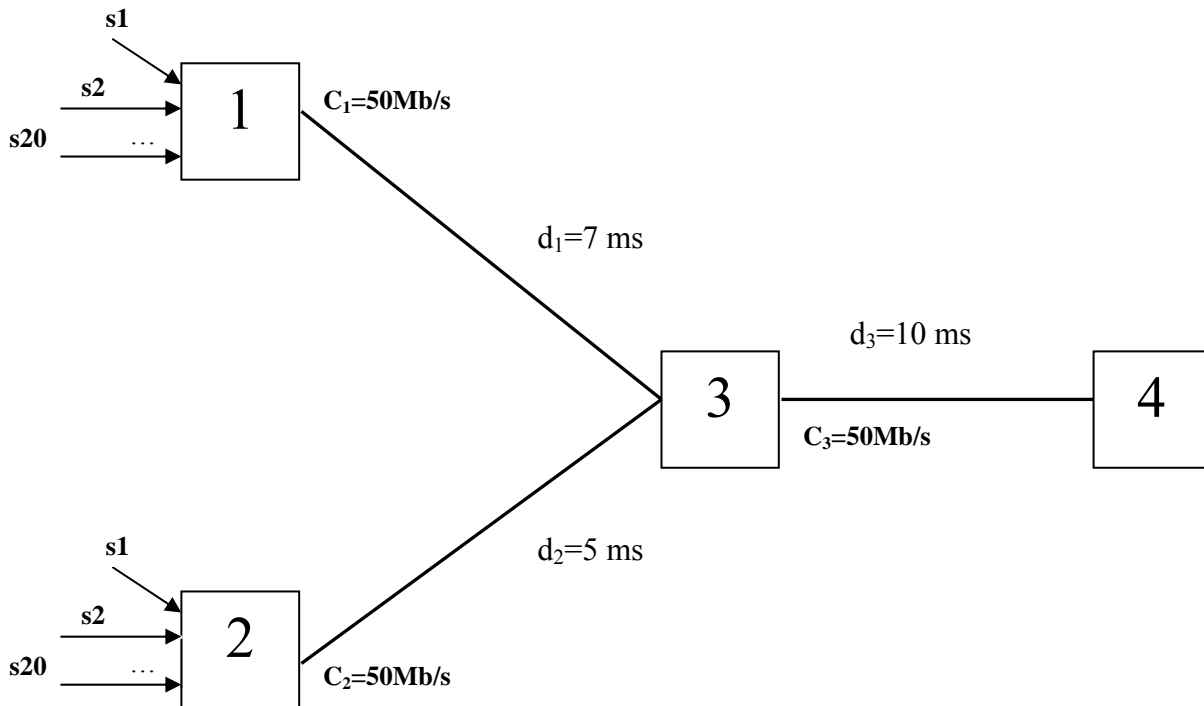


Figura 1: Topologia della rete da dimensionare.

Ad ogni edge router afferiscono 20 sorgenti TCP di tipo greedy. Il Core router, gestito con tecnica AQM di tipo RIO, è effettivamente il collo di bottiglia del sistema a causa della concorrenza che si instaura per il link a valle. La coda di uscita al nodo 3 tenderà infatti a saturarsi e l'AQM, scartando pacchetti accodati, frenerà le sorgenti impedendo che gli altri router scartino pacchetti. Tenendo conto di tale fatto è possibile semplificare le equazioni del modello trascurando la probabilità di perdita nelle code DropTail degli edge router (ciò equivale a dire che i buffer ai nodi 1 e 2 non sono mai pieni). Vedremo più avanti come i risultati delle simulazioni confermino tale ipotesi.

Il nostro obiettivo è garantire un Throughput minimo di 150 pacchetti/s per ognuna delle sorgenti a bassa priorità (out of profile) e di massimizzare quello delle sorgenti ad alta priorità mantenendo una probabilità di perdita sufficientemente bassa (minore dell'1% per traffico ad alta priorità e dell'8% per quello a bassa priorità). Considerando la condivisione del link 3-4 fra i due gruppi di flussi, ciò equivale a garantire una quota del 24% della capacità del sistema al traffico di bassa priorità e il rimanente per il traffico ad alta priorità.

RIO (RED with IN and OUT)

La Differentiated Services Architecture permette alle reti IP di offrire differenti livelli di QoS a differenti utenti ed applicazioni. In questa architettura i routers del core network offrono la stessa Per-Hop Behavior (PHB) a tutti i pacchetti classificati in una determinata classe negli edge router. Due dei più importanti tipi di PHB sono :

- Assured Forwarding (AF);
- Expedited Forwarding (EF);
- Best Effort (BE).

Per ogni tipo di PHB, sia esso AF che EF, i pacchetti IP sono marcati con differenti valori di drop precedence, e trattati differientemente nei nodi DS congestionati. A questo scopo i nodi DS nel core network implementano dei meccanismi di Active Queue Management (AQM).

Negli edge routers il token bucket marking mechanism classifica i pacchetti in due differenti drop precedence; una per i pacchetti che rispettano il profilo concordato (IN packets) e l'altra per i pacchetti che non lo rispettano (OUT packets).

Il **RIO** è un meccanismo implementato nei core routers allo scopo di trattare differientemente gli IN packets dagli OUT packets, in presenza di congestione, assegnando loro due diverse maschere RED; da qui il nome **RIO**, cioè RED with IN and OUT.

Metodologie utilizzate

Per risolvere il sistema di equazioni differenziali, abbiamo utilizzato un metodo di risoluzione alle differenze finite, approssimando le derivate con i loro rapporti incrementali ottenendo un set di equazioni alle differenze.

Utilizzando il linguaggio C in ambiente *linux*, abbiamo sviluppato un programma (allegato alla tesina) che consente di ottenere l'andamento temporale delle variabili di interesse.

Le equazioni del modello, particolareggiate per la topologia di rete da analizzare sono mostrate di seguito:

- equazioni della dinamica delle code nei router

$$\frac{dq_1(t)}{dt} = \begin{cases} -C_1 + \lambda_1^{(Q)}(t) & \text{se } q_1(t) > 0 \\ [-C_1 + \lambda_1^{(Q)}(t)]^+ & \text{se } q_1(t) = 0 \end{cases}$$

$$\frac{dq_2(t)}{dt} = \begin{cases} -C_2 + \lambda_2^{(Q)}(t) & \text{se } q_2(t) > 0 \\ [-C_2 + \lambda_2^{(Q)}(t)]^+ & \text{se } q_2(t) = 0 \end{cases}$$

$$\frac{dq_3(t)}{dt} = \begin{cases} -C_3 + [1 - p_{3-1}^{(B)}(t)] \cdot \lambda_1^{(Q)}(t) + [1 - p_{3-2}^{(B)}(t)] \cdot \lambda_2^{(Q)}(t) & \text{se } q_3(t) > 0 \\ [-C_3 + [1 - p_{3-1}^{(B)}(t)] \cdot \lambda_1^{(Q)}(t) + [1 - p_{3-2}^{(B)}(t)] \cdot \lambda_2^{(Q)}(t)]^+ & \text{se } q_3(t) = 0 \end{cases}$$

- equazione della coda stimata per il RED

$$\frac{dm_3(t)}{dt} = \begin{cases} \ln(1-\alpha) \cdot [m_3(t) - q_3(t)] \cdot \lambda_3^{(Q)}(t) & \text{se } q_3(t) > 0 \\ \ln(1-\alpha) \cdot C_3 \cdot m_3(t) & \text{se } q_3(t) = 0 \end{cases}$$

- equazioni del Round Trip Time

$$R_1(t) = 2d_1 + \frac{q_1(t)}{C_1} + 2d_3 + \frac{q_3(t)}{C_3}$$

$$R_2(t) = 2d_2 + \frac{q_2(t)}{C_2} + 2d_3 + \frac{q_3(t)}{C_3}$$

- equazioni della Transmission Window

$$\frac{dW_1(t)}{dt} \cong p_{3-1}^{(p)}(\tau_1) \frac{W_1(\tau_1)}{R_1(\tau_1)} \left[\left(1 - \frac{W_1(t)}{2}\right) Q(W_1(t)) - \frac{W_1(t)}{2} \right] + \begin{cases} \frac{(1 - p_{3-1}^{(p)}(\tau_1))W_1(\tau_1)}{R_1(\tau_1)} & \text{per } W_1(t) < T_1(t) \\ \frac{(1 - p_{3-1}^{(p)}(\tau_1))W_1(\tau_1)}{W_1(t)R_1(\tau_1)} & \text{per } W_1(t) \geq T_1(t) \end{cases}$$

$$\frac{dW_2(t)}{dt} \cong p_{3-2}^{(p)}(\tau_2) \frac{W_2(\tau_2)}{R_2(\tau_2)} \left[\left(1 - \frac{W_2(t)}{2}\right) Q(W_2(t)) - \frac{W_2(t)}{2} \right] + \begin{cases} \frac{(1 - p_{3-2}^{(p)}(\tau_2))W_2(\tau_2)}{R_2(\tau_2)} & \text{per } W_2(t) < T_2(t) \\ \frac{(1 - p_{3-2}^{(p)}(\tau_2))W_2(\tau_2)}{W_2(t)R_2(\tau_2)} & \text{per } W_2(t) \geq T_2(t) \end{cases}$$

- equazioni della Treshold

$$\frac{dT_1(t)}{dt} \cong p_{3-1}^{(p)}(\tau_1) \frac{W_1(\tau_1)}{R_1(\tau_1)} \left[\frac{W_1(t)}{2} - T_1(t) \right]$$

$$\frac{dT_2(t)}{dt} \cong p_{3-2}^{(p)}(\tau_2) \frac{W_2(\tau_2)}{R_2(\tau_2)} \left[\frac{W_2(t)}{2} - T_2(t) \right]$$

- equazioni della Probabilità di perdita sui path (coincidente, nel nostro caso, con la probabilità di perdita al nodo 3):

$$p_1^{(P)}(t) = p_{3-1}^{(B)}(t) = \begin{cases} 0 & \text{se } m_3(t) < t_{\min 1} \\ \frac{m_3(t) - t_{\min 1}}{t_{\max 1} - t_{\min 1}} \cdot p_{\max 1} & \text{se } t_{\min 1} < m_3(t) < t_{\max 1} \\ 1 & \text{se } m_3(t) > t_{\max 1} \end{cases}$$

$$p_2^{(P)}(t) = p_{3_2}^{(B)}(t) = \begin{cases} 0 & \text{se } m_3(t) < t_{\min 2} \\ \frac{m_3(t) - t_{\min 2}}{t_{\max 2} - t_{\min 2}} \cdot p_{\max 2} & \text{se } t_{\min 2} < m_3(t) < t_{\max 2} \\ 1 & \text{se } m_3(t) > t_{\max 2} \end{cases}$$

- equazioni del Throughput

$$Th_1(t) = \frac{W_1(t)}{R_1(t)} [1 - p_1^{(P)}(t)]$$

$$Th_2(t) = \frac{W_2(t)}{R_2(t)} [1 - p_2^{(P)}(t)]$$

Tali equazioni possono essere trasformate agevolmente in equazioni alle differenze; aggiungendo le condizioni iniziali:

$$q_1(0) = q_2(0) = q_3(0) = 0 \text{ pacchetti}$$

$$m_3(0) = 0 \text{ pacchetti}$$

$$R_1(0) = 2d_1 + 2d_3 = 34 \text{ ms}$$

$$R_2(0) = 2d_2 + 2d_3 = 30 \text{ ms}$$

$$W_1(0) = W_2(0) = 1 \text{ pacchetti}$$

$$T_1(0) = T_2(0) = 128 \text{ pacchetti}$$

$$p_{3_1}^{(B)}(0) = p_{3_2}^{(B)}(0) = 0 \%$$

$$Th_1(0) = \frac{W_1(0)}{R_1(0)} = \frac{1}{34 \cdot 10^{-3}} = 29.41 \text{ pacchetti/s}$$

$$Th_2(0) = \frac{W_2(0)}{R_2(0)} = \frac{1}{30 \cdot 10^{-3}} = 33.33 \text{ pacchetti/s}$$

si perviene a una soluzione numerica rappresentata da un insieme di vettori contenenti i campioni delle soluzioni, campionate con un certo tempo di campionamento.

Avendo osservato, dai primi risultati ottenuti, che il sistema andava a regime già dopo circa 2s, abbiamo fissato la durata della finestra di osservazione a 20s e il tempo di campionamento pari a 1ms, per poter avere un maggior numero di campioni da utilizzare per il calcolo delle medie.

Per ricavare dei valori rappresentativi della singola analisi effettuata si eseguono infine le medie aritmetiche dei vettori rappresentanti il Throughput e la probabilità di perdita (che naturalmente escludono i primi 3 secondi di transitorio).

Allo scopo di semplificare la progettazione abbiamo deciso di fissare a priori i parametri della maschera RED per il flusso a minore priorità. Assegniamo a tali parametri i seguenti valori:

- $t_{\min 2} = 50$ pacchetti
- $t_{\max 2} = 80$ pacchetti
- $p_{\max 2} = 30\%$

Le figure 2 e 3 mostrano l'andamento temporale del throughput per due diversi valori di α :

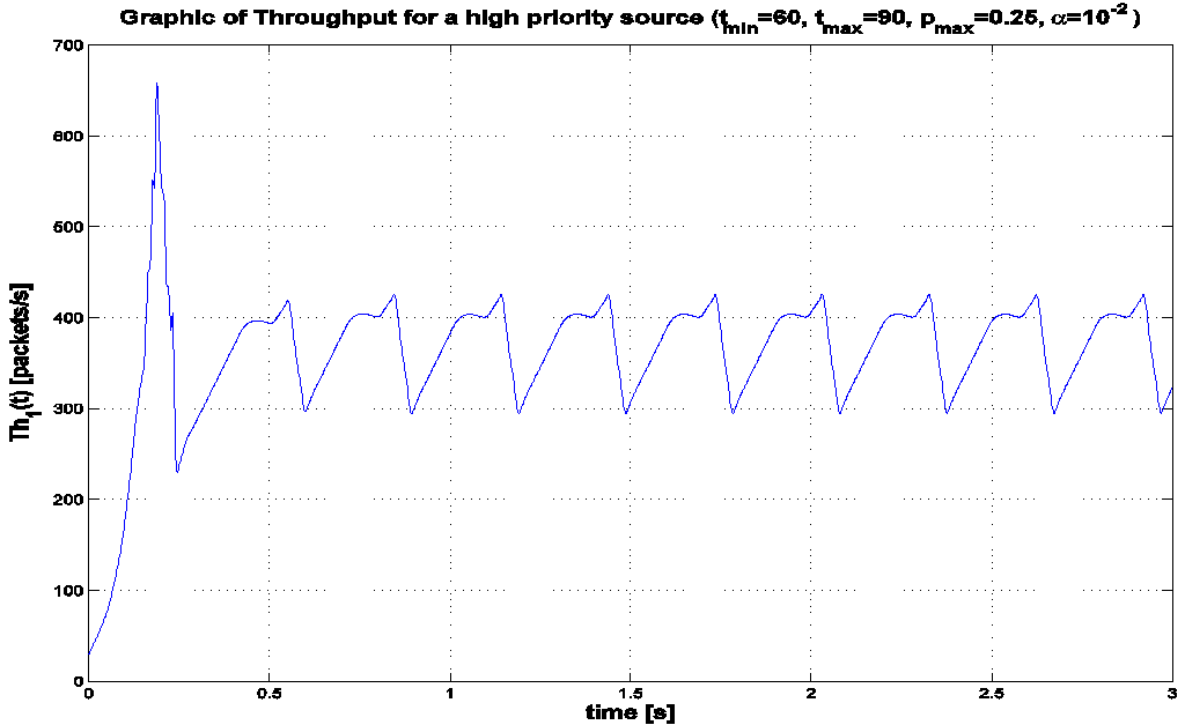


Figura 2: Andamento del throughput per il primo gruppo di sorgenti per $\alpha=10^{-3}$.

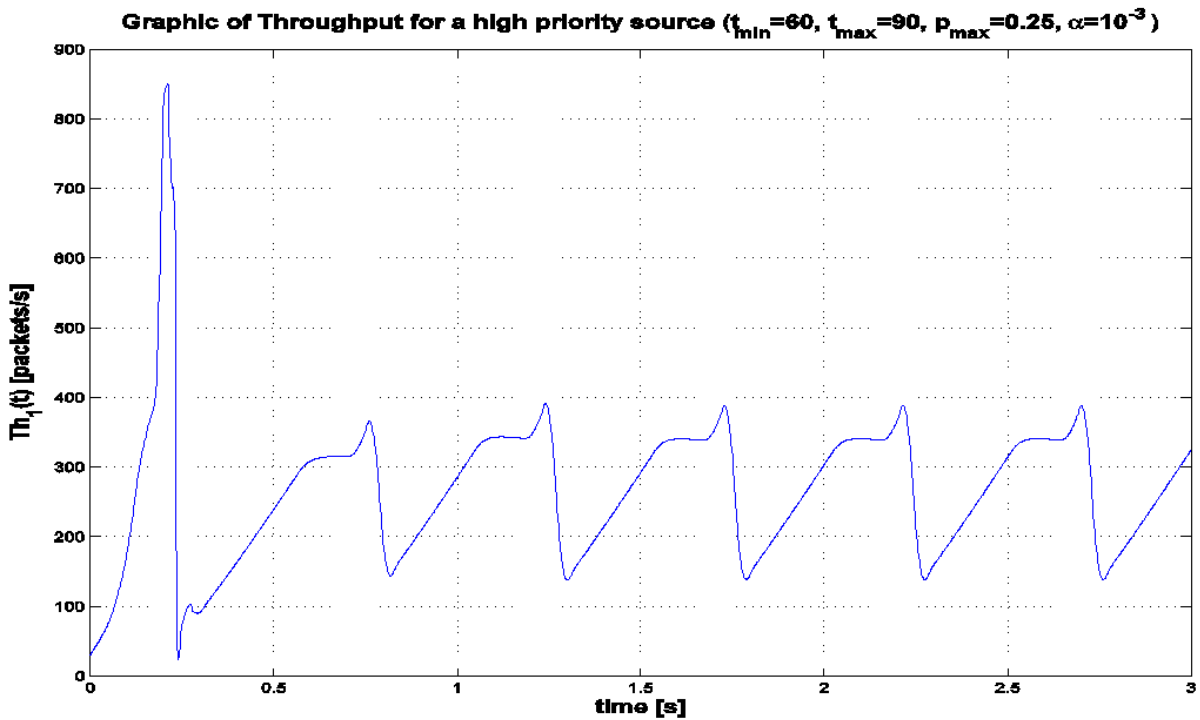


Figura 3: Andamento del throughput per il primo gruppo di sorgenti per $\alpha=10^{-2}$.

Dai Grafici si vede che il throughput aumenta per valori di α più grandi, di contro si ha una più evidente oscillazione nel transitorio della coda del router RIO (vedi figura 4). È stato scelto un

valore di α pari a 10^{-2} perché dalle prime prove si è trovato un buon compromesso fra stabilità e throughput.

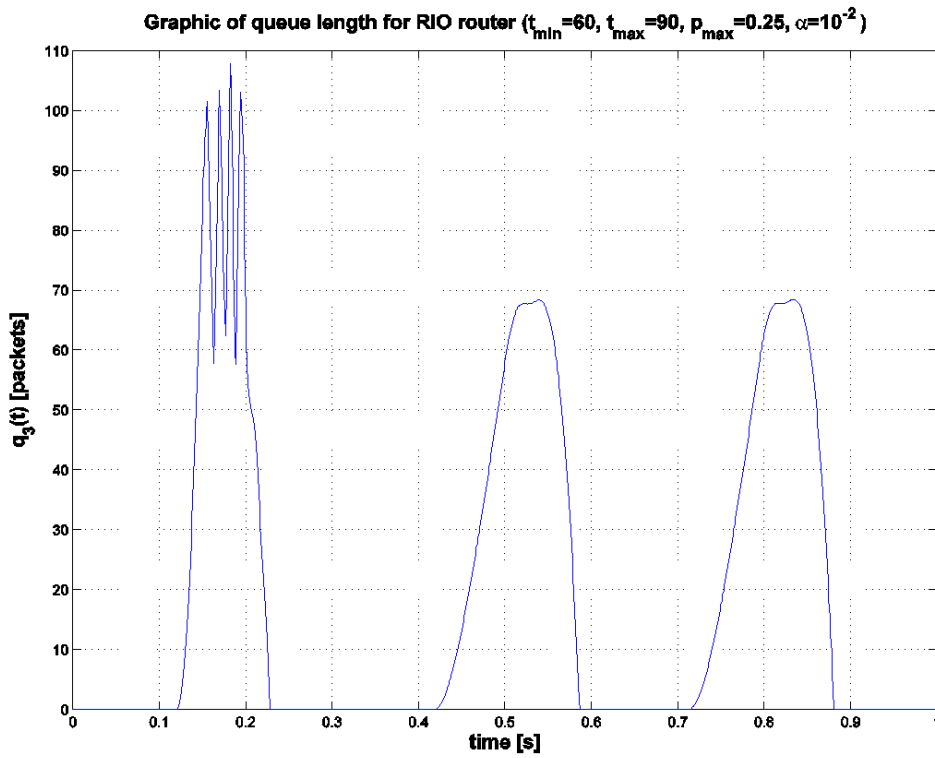


Figura 4: Andamento della lunghezza della coda al nodo 3 per il primo gruppo di sorgenti con $\alpha=10^{-2}$.

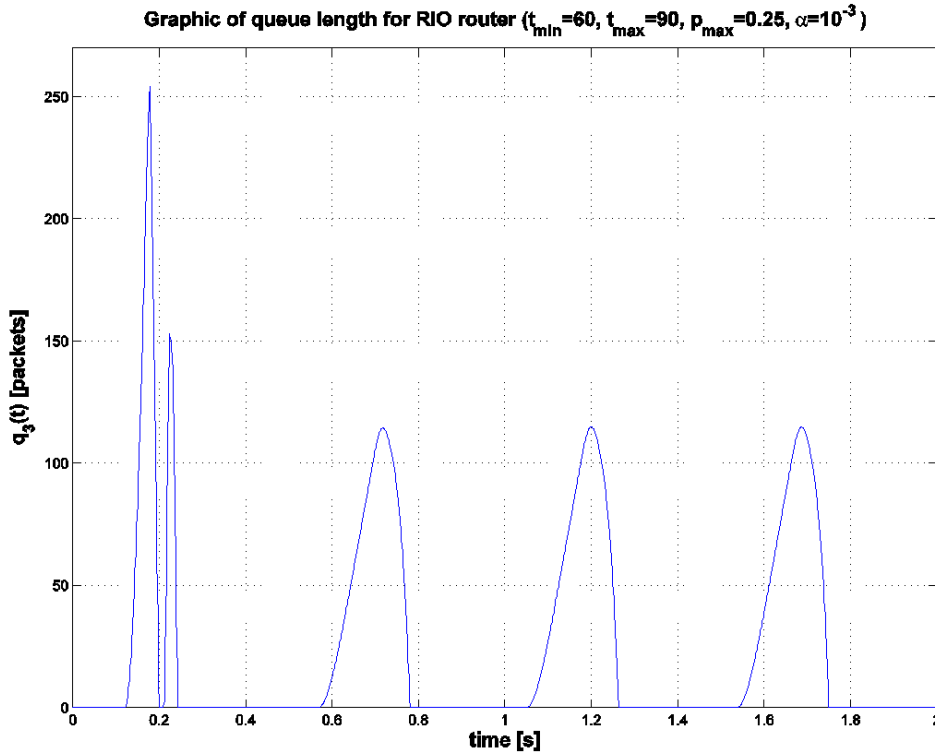


Figura 5: Andamento della lunghezza della coda al nodo 3 per il primo gruppo di sorgenti con $\alpha=10^{-3}$.

Risultati ottenuti

Allo scopo di trovare le configurazioni che garantiscano le prestazioni migliori abbiamo graficato (figura 6) il Throughput e la probabilità di perdita medi relativi al flusso 1 (quello a maggior priorità), in funzione del parametro t_{min1} per diversi valori della differenza $\Delta = t_{max1} - t_{min1}$.

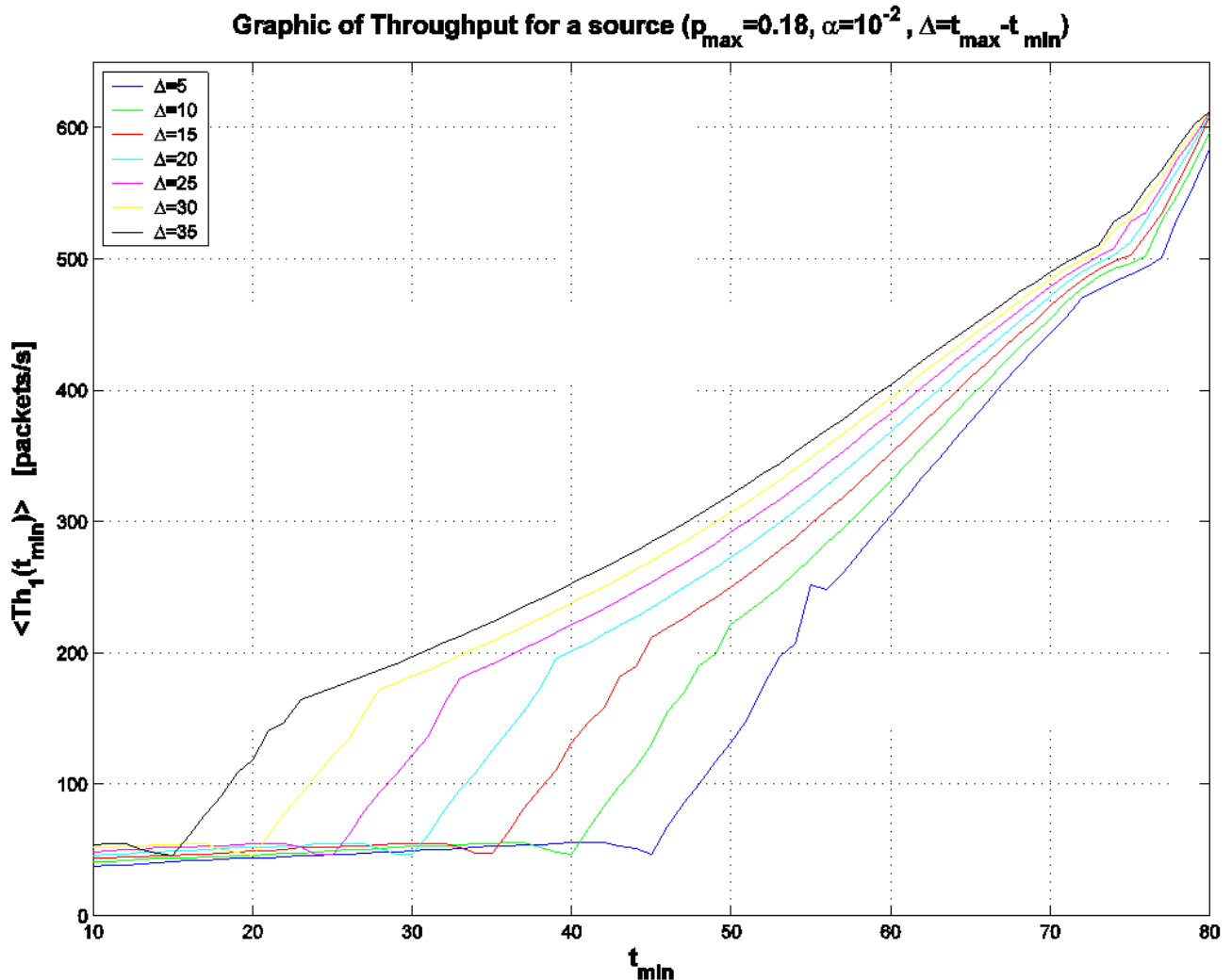


Figura 6: Andamento del Throughput medio al variare di t_{min1} per diversi valori della differenza Δ .

Dall'osservazione della figura 6 si nota come, la scelta del valore di t_{min1} sia decisiva per avere un throughput sufficientemente alto. Infatti considerando, ad esempio, la curva per $\Delta=20$ si vede che al di sotto di t_{min1} pari a 30 il throughput è estremamente basso, mentre per t_{min1} superiore a 45 si ha un throughput accettabile e con andamento crescente all'aumentare di t_{min1} .

Dalla figura 7, in cui sono graficati gli andamenti della probabilità di perdita media al variare di t_{min1} per diversi valori della differenza Δ , si nota come le precedenti considerazioni trovano riscontro anche per quanto riguarda la probabilità di perdita. Infatti è necessario porsi alla destra del ginocchio per avere probabilità di perdita accettabili; ad esempio, per $\Delta=20$ e $t_{min1}>45$ si ha una probabilità di perdita inferiore al 3%.

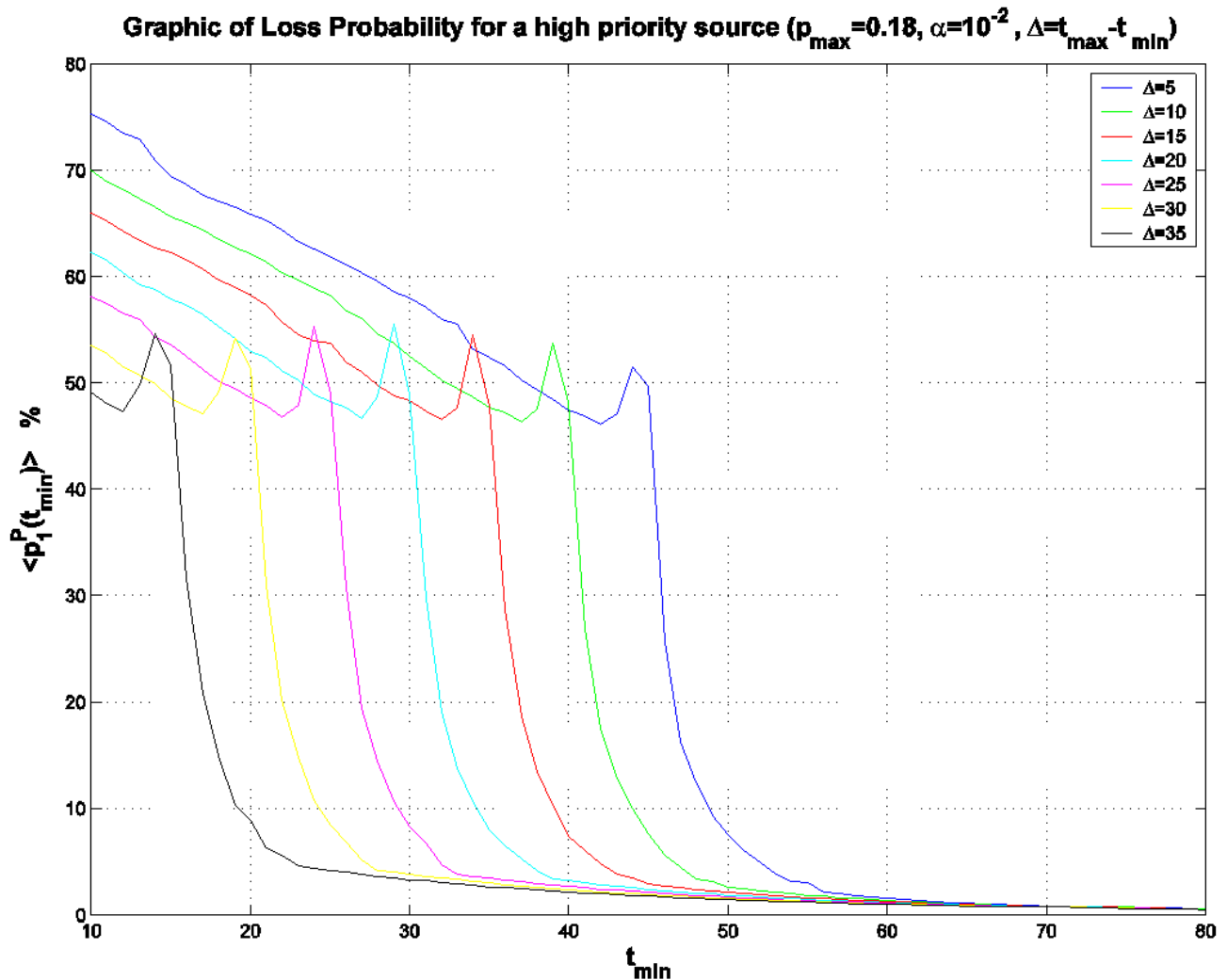


Figura 7: Andamento della probabilità di perdita media al variare di t_{min1} per diversi valori della differenza Δ .

Le scelte di trade-off ci portano a considerare ottimale la curva con $\Delta=20$ e scegliere un valore di t_{min1} superiore a 45.

La scelta di p_{max1} non è risultata determinante ai fini del dimensionamento prefissato; sostanzialmente al crescere di p_{max1} ed a parità di altri parametri si ha una diminuzione delle prestazioni del flusso a maggiore priorità (riduzione del throughput e aumento della probabilità di perdita) a vantaggio di quelle a bassa priorità. Per ottenere il giusto equilibrio fra i due tipi di flusso, abbiamo scelto un valore per p_{max1} pari a 0.18.

Muovendoci lungo il tratto di curva crescente in figura 6, e controllando anche i valori medi dei throughput per le sorgenti a bassa priorità, otteniamo una serie di valori rappresentati in tabella 1.

t_{min1}	t_{max1}	$\overline{Th}_1^{(P)}$	$\overline{Th}_2^{(P)}$	$\overline{p}_1^{(P)}$	$\overline{p}_2^{(P)}$
50	70	272.96	293.78	1.85	1.83
60	80	368.53	219.37	1.13	3.21
65	85	421.50	183.80	0.92	4.49
70	90	472.22	151.19	0.79	6.48
75	95	512.45	123.78	0.67	9.11

Tabella 1: Valori medi ottenuti delle grandezze di interesse al variare di t_{min1} per $\Delta=20$.

Dall'esame della tabella, si capisce che il valore di t_{min1} che ci consente di rispettare i requisiti sulle sorgenti a bassa priorità e di massimizzare il throughput di quelle ad alta priorità è 70.

In definitiva i valori dei parametri da dimensionare che soddisfano tutte le condizioni imposte e che ottimizzano le caratteristiche (throughput e probabilità di perdita) del flusso ad alta priorità sono:

$$t_{\min 1} = 70 \text{ pacchetti}$$

$$t_{\max 1} = 90 \text{ pacchetti}$$

$$p_{\max 1} = 18 \%$$

$$\alpha = 10^{-2}$$

$$t_{\min 2} = 50 \text{ pacchetti}$$

$$t_{\max 2} = 80 \text{ pacchetti}$$

$$p_{\max 2} = 30 \%$$

Di seguito sono mostrate le forme d'onda in funzione del tempo per tutte le variabili di interesse ottenute utilizzando i parametri precedentemente elencati.

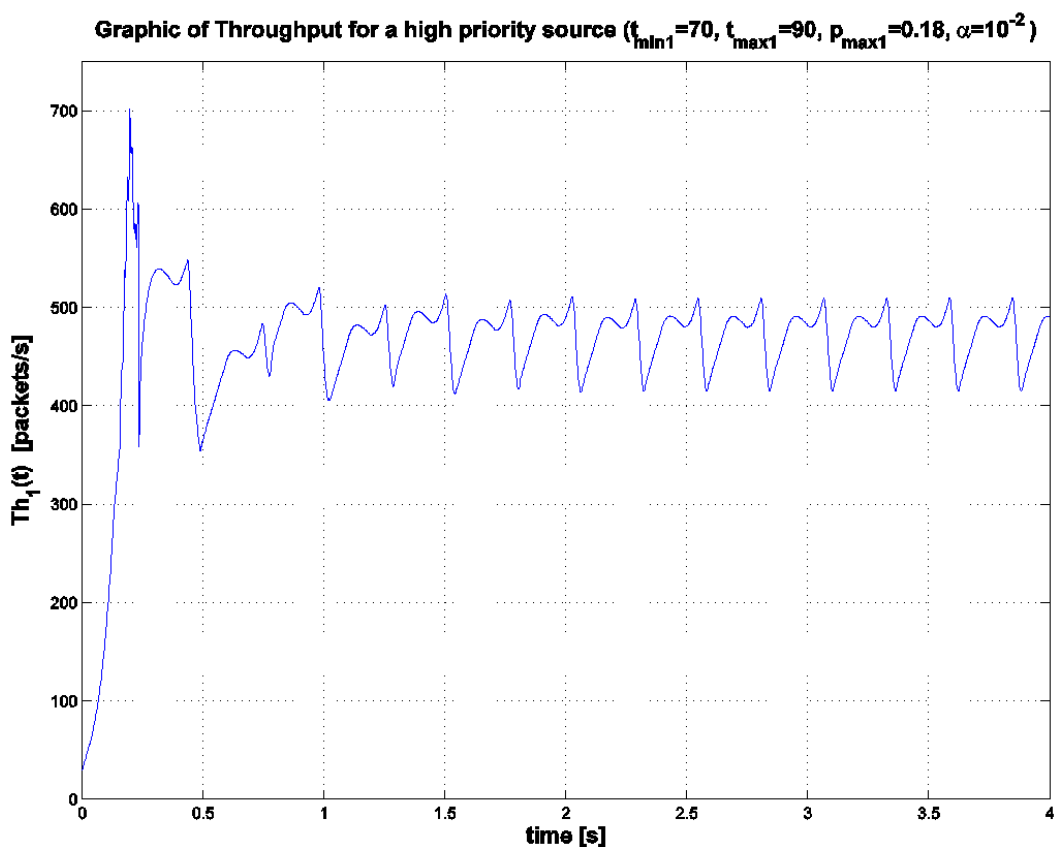


Figura 8: Andamento del throughput della generica sorgente ad alta priorità.

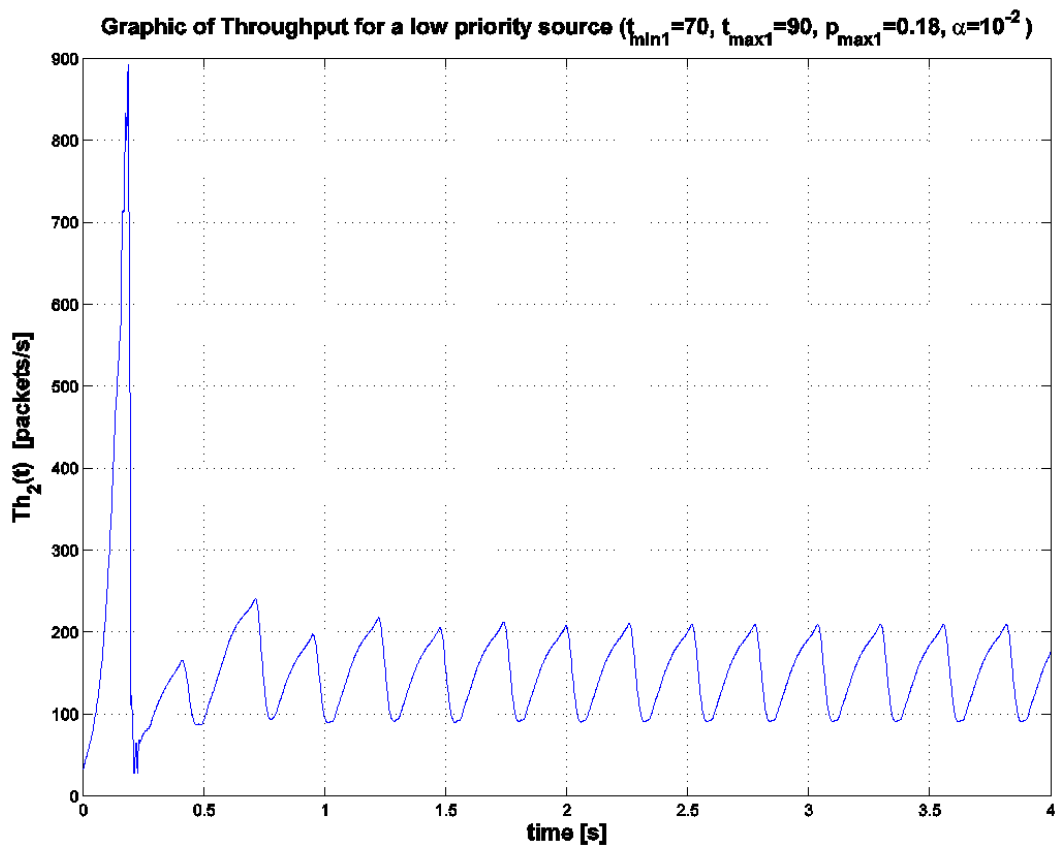


Figura 9: Andamento del throughput della generica sorgente a bassa priorità.

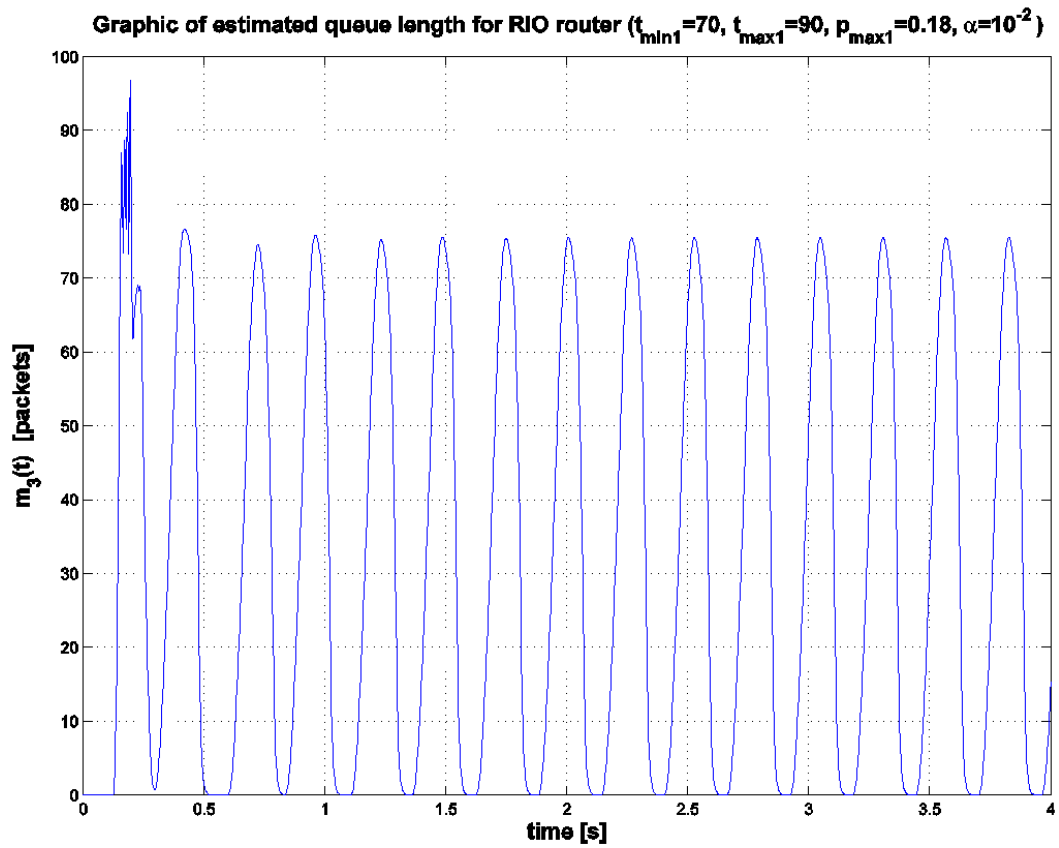


Figura 10: Andamento della coda stimata per il router RIO.

Graphic of Congestion Window for the first group of sources ($t_{\min 1}=70, t_{\max 1}=90, p_{\max 1}=0.18, \alpha=10^{-2}$)

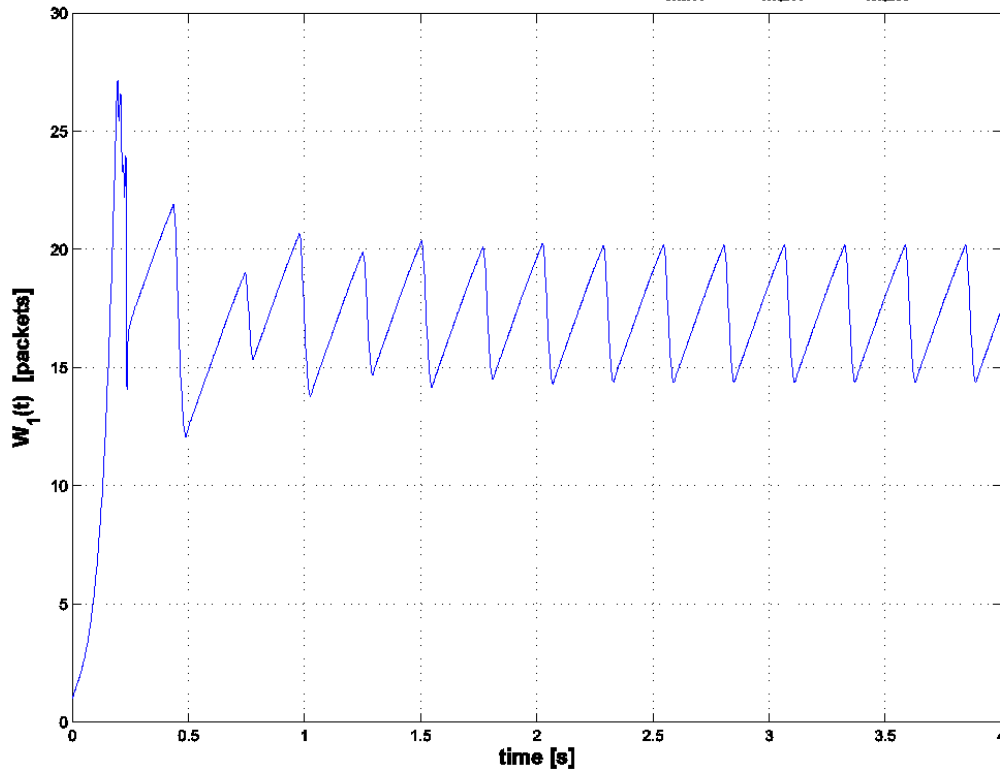


Figura 11: Andamento della Congestion Window per la generica sorgente ad alta priorit .

Graphic of Congestion Window for the second group of sources ($t_{\min 1}=70, t_{\max 1}=90, p_{\max 1}=0.18, \alpha=10^{-2}$)

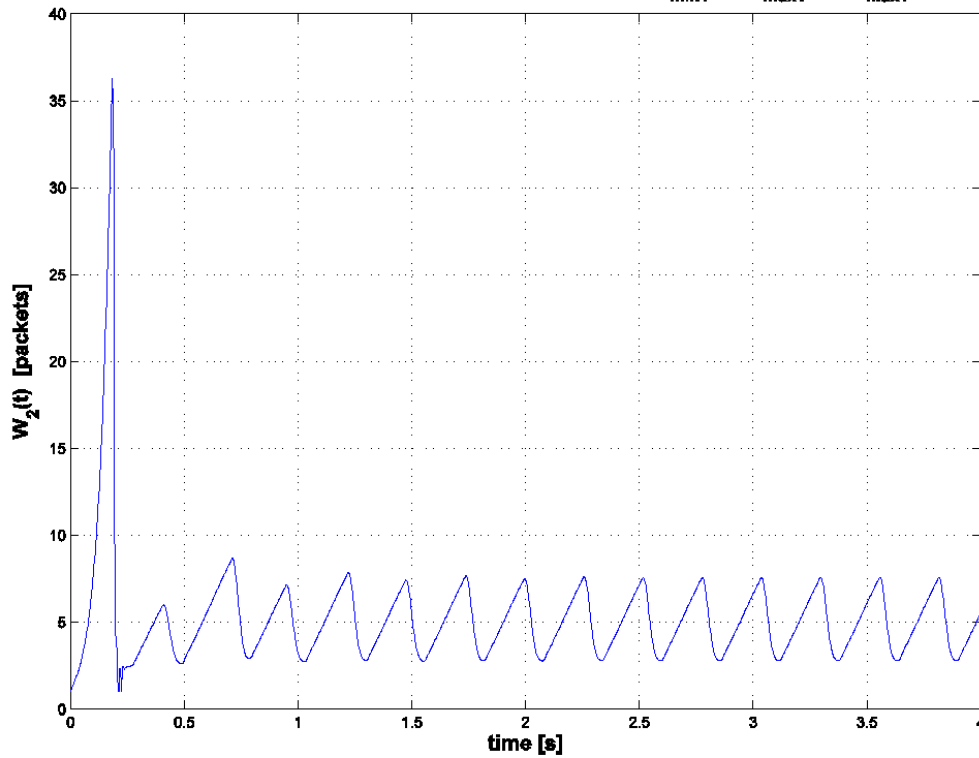


Figura 12: Andamento della Congestion Window per la generica sorgente a bassa priorit .

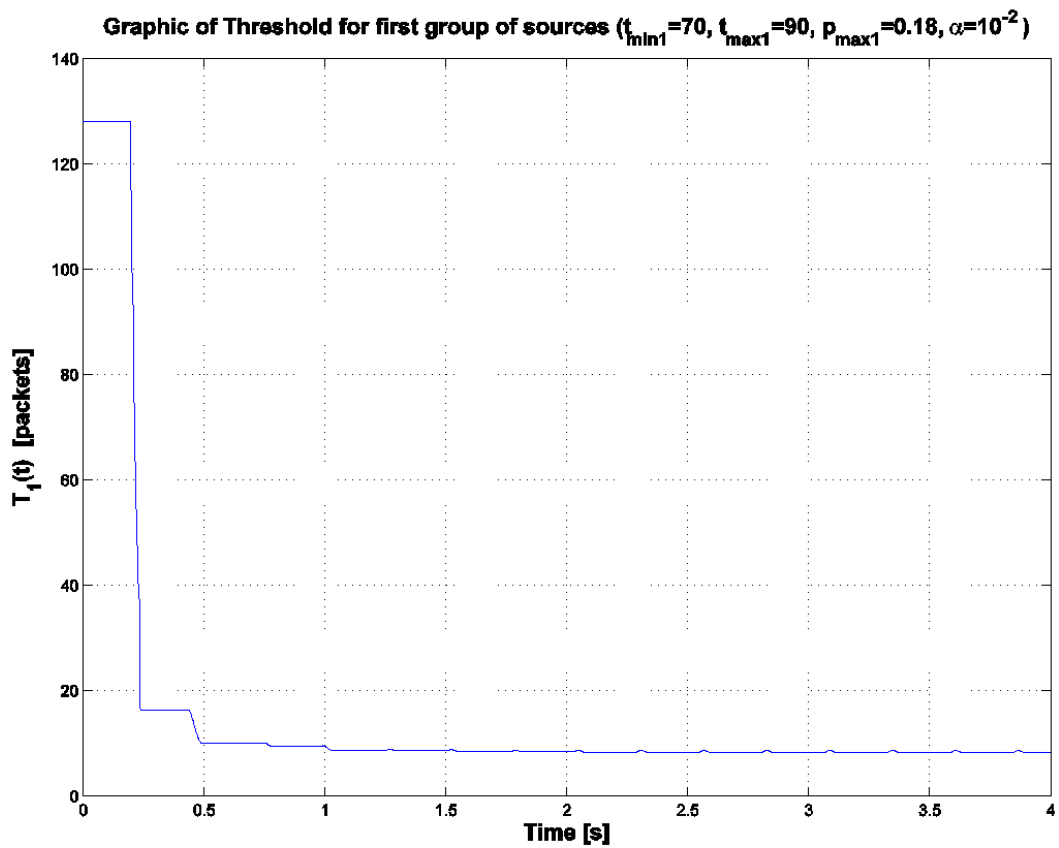


Figura 13: Andamento della Threshold per la generica sorgente ad alta priorit .

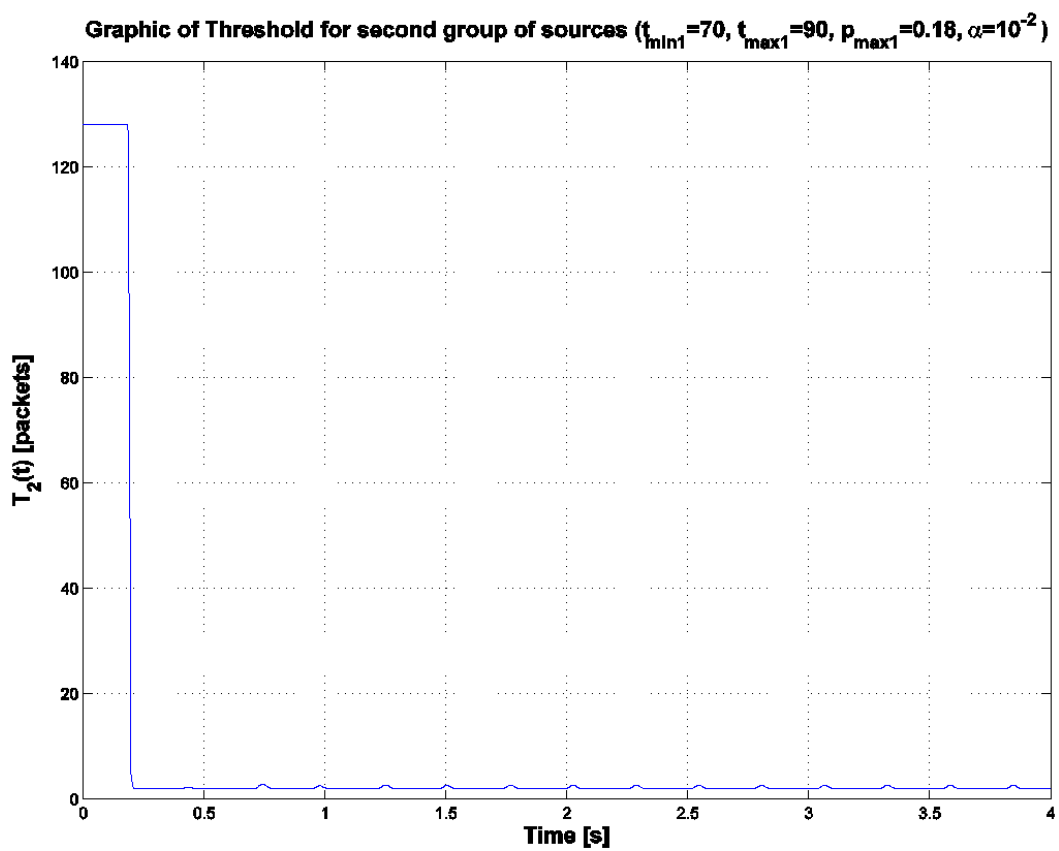


Figura 14: Andamento della Threshold per la generica sorgente a bassa priorit .

Graphic of Round Trip Time for the first group of sources ($t_{\min 1}=70, t_{\max 1}=90, p_{\max 1}=0.18, \alpha=10^{-2}$)

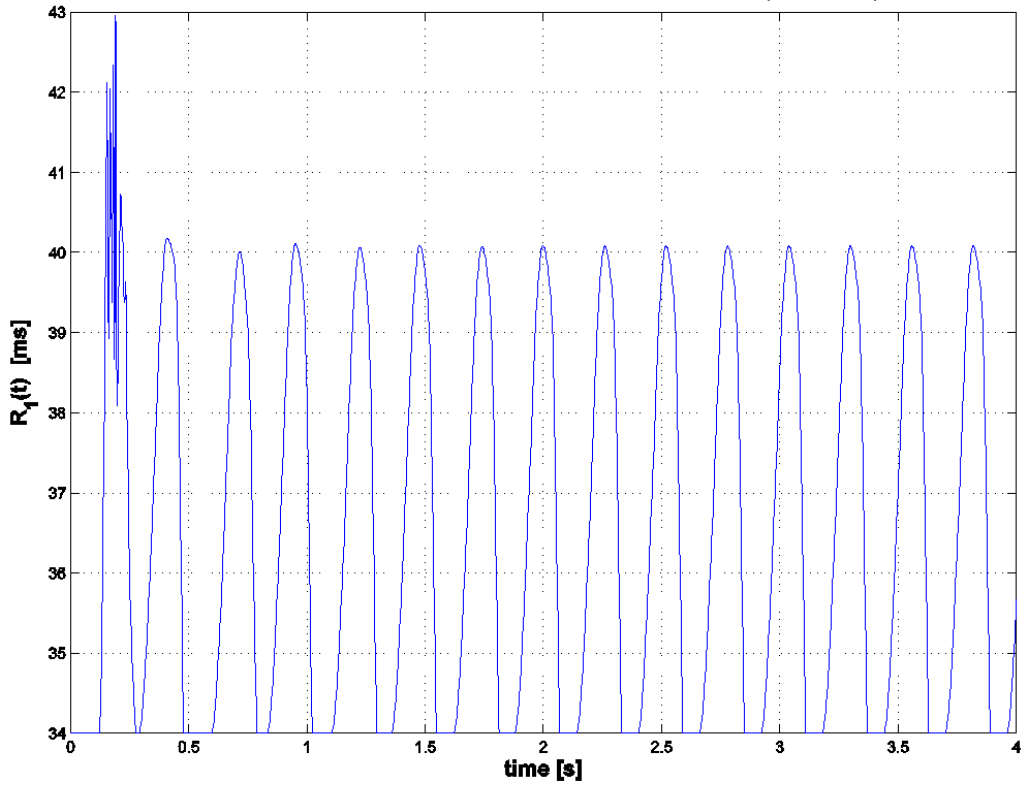


Figura 15: Andamento del Round Trip Time per la generica sorgente ad alta priorit .

Graphic of Round Trip Time for the second group of sources ($t_{\min 1}=70, t_{\max 1}=90, p_{\max 1}=0.18, \alpha=10^{-2}$)

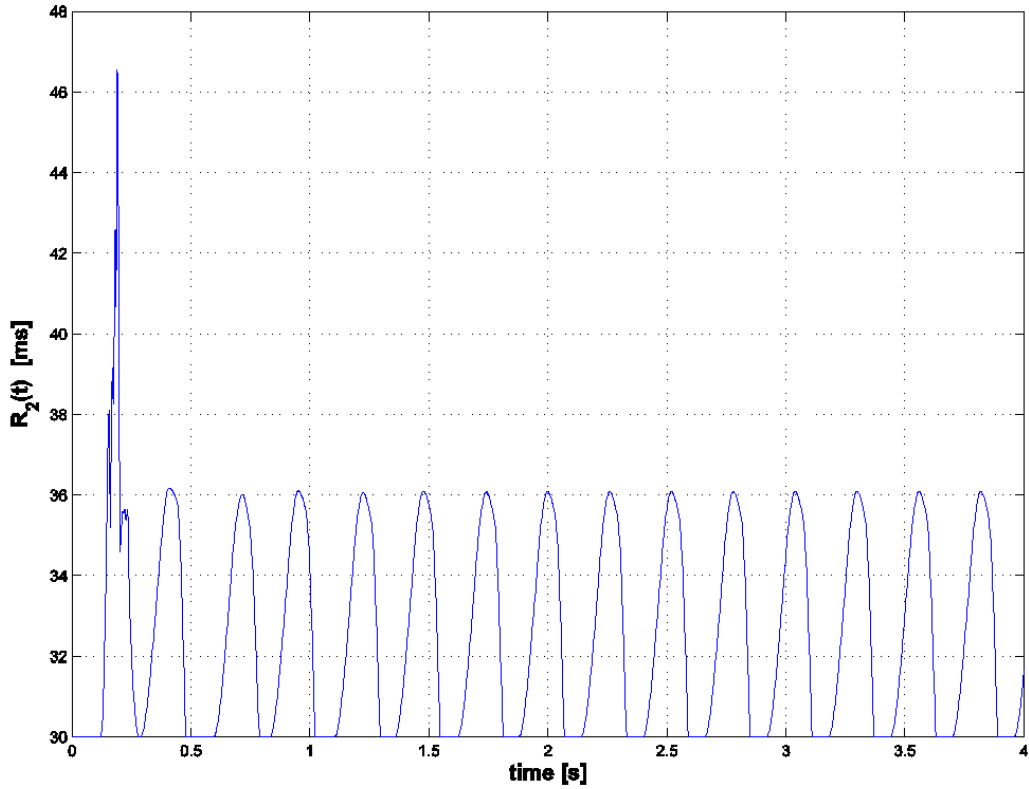


Figura 16: Andamento del Round Trip Time per la generica sorgente ad alta priorit .

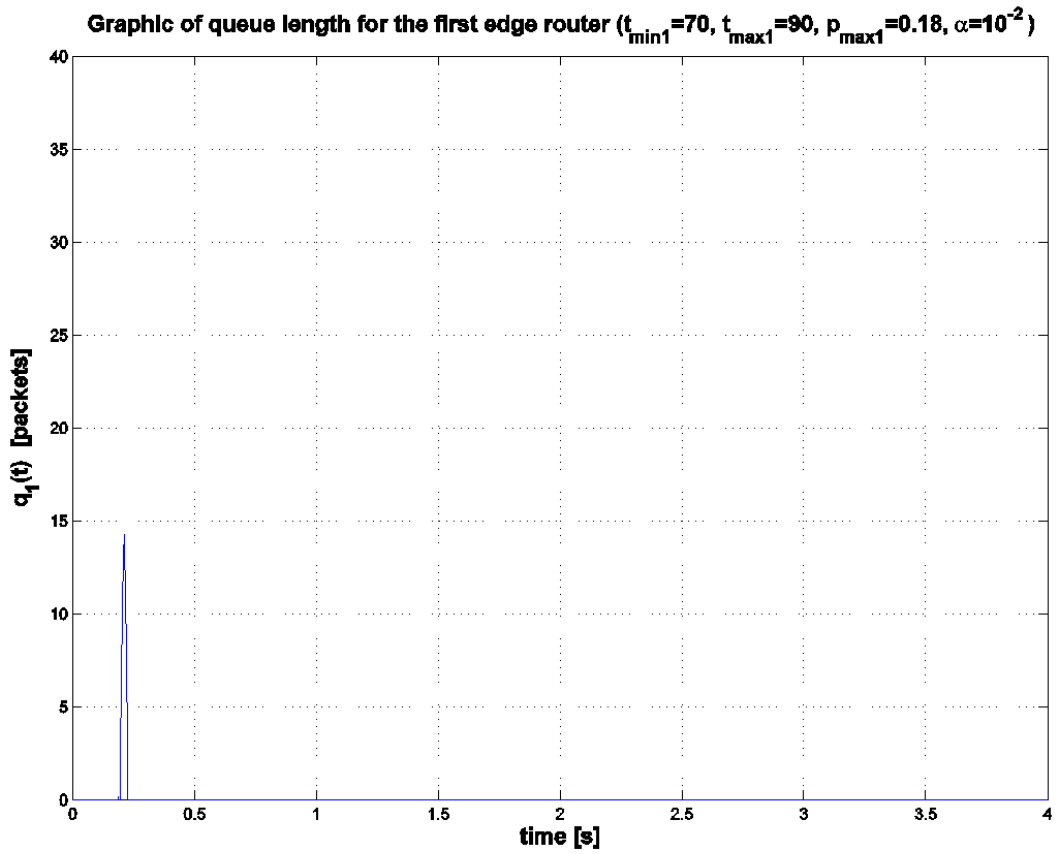


Figura 17: Andamento della coda del primo edge router.

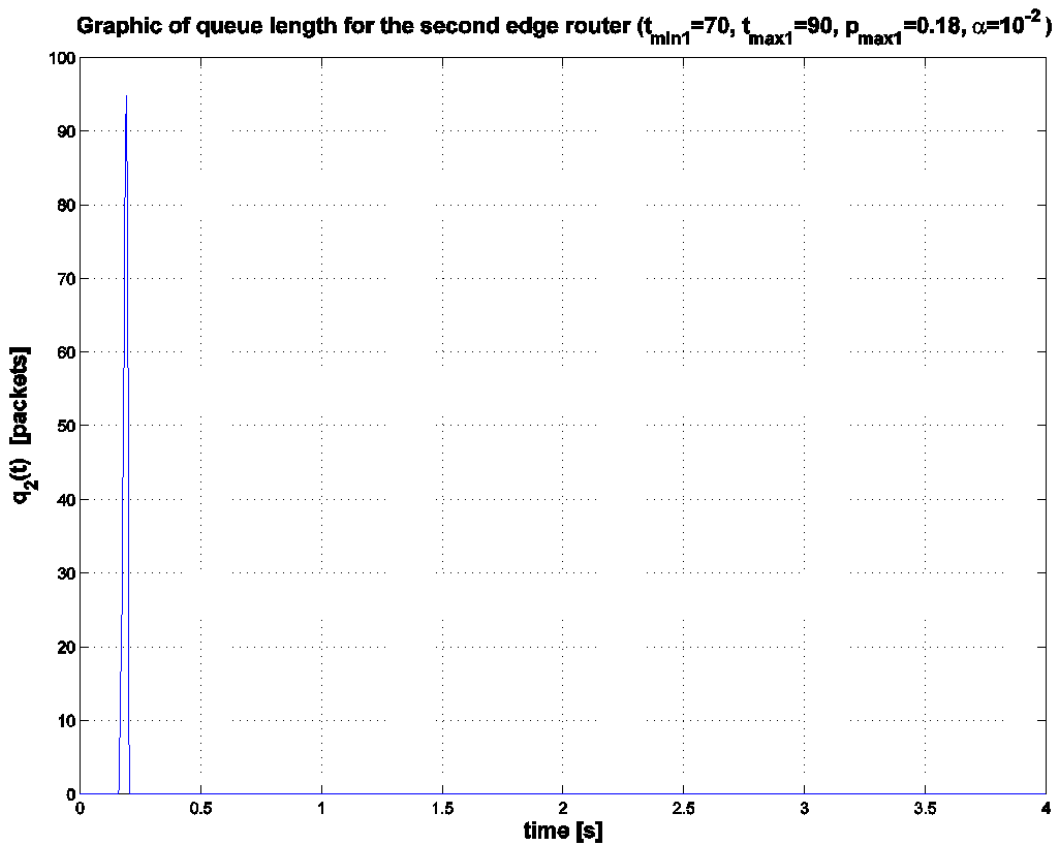


Figura 18: Andamento della coda del secondo edge router.

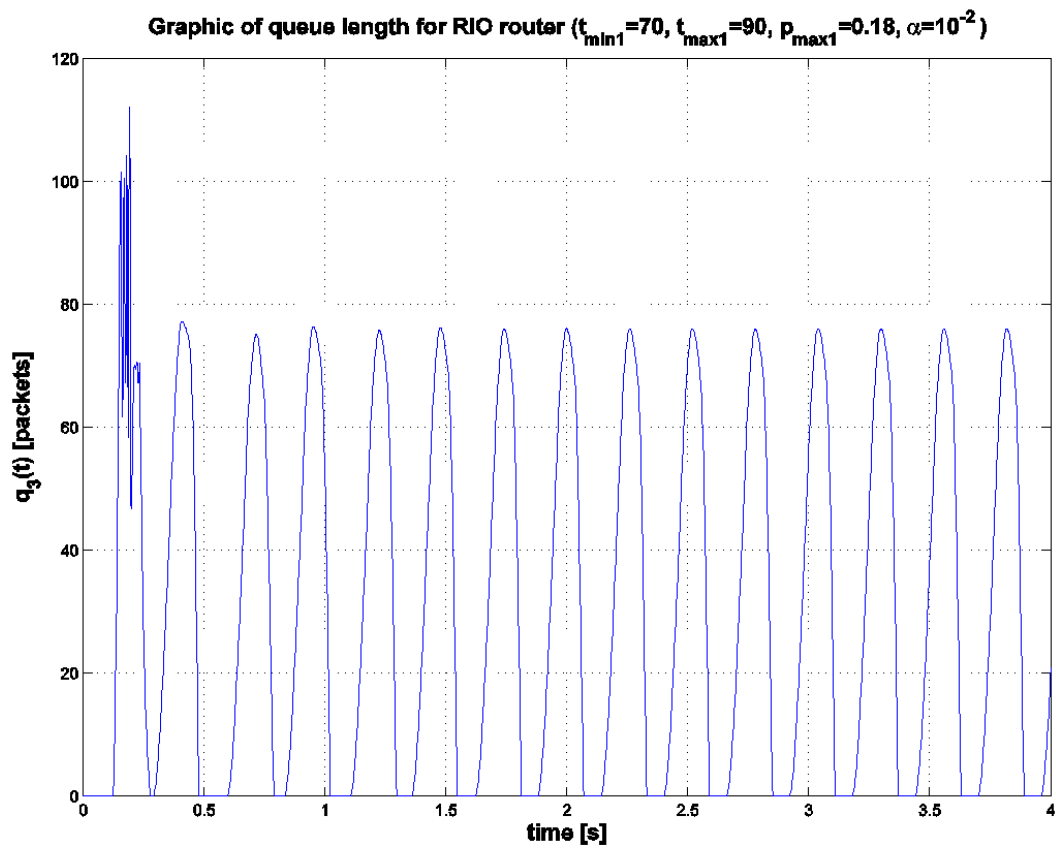


Figura 19: Andamento della coda del router RIO.