A Simple Formula for End-to-End Jitter Estimation in Packet-switching Networks

Olivier Brun, Charles Bockstal, Jean-Marie Garcia

Abstract— **In this paper, an analytic approximation is derived for the end-to-end delay-jitter incurred by a periodic traffic with constant packet size. It is assumed that the periodic traffic is multiplexed with a background packet stream under the FCFS service discipline in each queue along the path to its destination. The processes governing the packet arrivals and the packet sizes of the background traffics are assumed to be general renewal processes. A very simple analytical approximation is derived and its accuracy is assessed by means of event-driven simulations.**

Keywords— **delay-jitter, $GI+D/GI+D/1$ queue, Internet, CBR traffic**

I. Introduction

WITH more than 800 millions worldwide users, Internet has become pervasive in the last decade thanks to a succession of drivers, starting with the rise of the World Wide Web, followed by business adoption of Internet applications, and more recently, the use of peer-to-peer file-sharing softwares. The Internet is now expected to become a global communication infrastructure supporting real-time multimedia applications with stringent Quality-of-Service (QoS) requirements.

Many real-time speech or video applications are generating Constant Bit Rate (CBR) traffic streams. An important performance measure for such traffics is the jitter which can be defined as the distortion on the periodic nature of the packet stream between the source and the destination. This distortion of the original pattern is due to the series of multiplexing operations performed by network routers.

In the last decade, many studies have been devoted to the estimation of jitter in packet-switched networks. Most of these works have been done in the context of ATM networks and therefore assume discrete time models. Jitter has been analyzed by means of event-driven simulations [3], [5]. Bounds have also been derived, either in a deterministic setting [6], [7], [8], [9] or in a stochastic setting [10], [11], but they are considered too loose for practical purposes. Other previous works are based on queueing theory [12], [13], [14], [15], [16], [18], [19], [20], [21], [23], [24], [26], [27], [28]. Exact and approximate numerical methods have been developed for the single node case. In [20], the node-by-node approximation of jitter is used for the numerical evaluation of the end-to-end jitter. The main drawback of these numerical models is that computations are typically very expensive. To the extend of our knowledge, only [25] proposes an analytic formula of end-to-end jitter based on heavy and light traffic approximations. The range of validity of this formula is however limited.

Authors are with the LAAS-CNRS, Toulouse, France. E-mail : brun@laas.fr, jmg@laas.fr, cbocksta@laas.fr

In this paper, a simple analytic approximation is derived for the end-to-end delay-jitter incurred by a periodic traffic having a constant packet size. We assume that the periodic traffic is multiplexed with a background traffic under the FCFS service discipline. The only assumption regarding the background traffic is that the processes governing packet arrivals and packet sizes are renewal processes. The accuracy of the approximation is assessed by means of numerous event-driven simulations. In our opinion, having such simple and yet accurate formula for jitter estimation is fundamental if one considers the simulation of networks with thousands of flows or the iterative algorithms for network optimization subject to QoS constraints.

The paper is organized as follows. Our assumptions and notations are introduced in the following section. Section III introduces the exponential approximation for tail probabilities of the steady-state waiting time that will be used throughout the paper. Section IV is devoted to the single-node case and section V extends the approximation of the single-node jitter to the multiple-node case. Finaly, some conclusions are drawn in section VI.

II. Problem statement

We consider a tagged traffic passing through n tandem queues. This tagged traffic is assumed to be initialy periodic, i.e. we assume a constant time $T = 1/\lambda_T$ between packet arrivals at node 1. All the packets of the tagged traffic have the same size: D data units. The service rate of queue $k = 1 \dots n$ is μ_k data units per second.

In each queue $k = 1, \dots, n$, the tagged traffic is multiplexed under the FCFS discipline with a background traffic. Packet arrivals of the background traffic as well as packet sizes are assumed to be governed by general renewal processes, with respective probability density functions (pdf) $a_k^{\rm bg}(t)$ and $b_k^{\rm bg}(x)$. Let $\lambda_k^{\rm bg}$ denotes the packet rate of the background traffic at queue k and let $x_k^{\rm bg}$ denotes its average packet size (in data units). The global utilization factor of queue k is therefore $\rho_k = \rho_k^{\rm bg} + (\lambda_T D)/\mu_k$ where $\rho_k^{\rm bg} = \lambda_k^{\rm bg}\, x_k^{\rm bg}/\mu_k$ is the offered load of the background traffic. In the following, $\rho_k < 1$ is assumed for $k = 1, \dots, n$.

Let us consider two consecutive packets of the tagged traffic, C_0 and C_1. For packet C_j, $j = 0, 1$ and node $k = 1, \dots, n$, let:

- $\tau_j^{\text{in}}(k)$ and $\tau_j^{\text{out}}(k)$ be the arrival and departure times of C_j at node k. Without loss of generality, we assume that $\tau_0^{\text{in}}(1) = 0$,
- $S_j(k) = \tau_j^{\text{out}}(k) - \tau_j^{\text{in}}(k)$ be the sojourn time of C_j at node k,
- $W_j(k) = S_j(k) - D/\mu_k$ be the waiting time of C_j at node k,
- $\Delta_k = W_1(k) - W_0(k)$ be the variation of the inter-packet delay at node k,

End-to-end jitter is defined as,

$$J_{[1\ldots n]}(T) = E\left[\ |(\tau_1^{\text{out}}(n) - \tau_0^{\text{out}}(n)) \ - \ T\ |\ \right]$$

Since $\tau_i^{\text{in}}(k+1) = \tau_i^{\text{out}}(k)$ for all $i = 0, 1$ and $k = 1, \ldots, n-1$, it is easy to show that,

$$\begin{aligned} J_{[1\ldots n]}(T) &= E\left[\ |\sum_{k=1}^{n}(W_1(k) - W_0(k))\ |\ \right] \\ &= E\left[\ |\sum_{k=1}^{n}\Delta_k\ |\ \right] \end{aligned} \tag{1}$$

The end-to-end jitter is therefore given by the expected absolute value of the sum of inter-packet delay variations introduced by each node along the path between the source and the destination. In the following, we will derive an approximation of the probability density function $f_{\Delta_1+\ldots+\Delta_n}(x)$ of this sum.

III. Exponential approximation for tail probabilities of the waiting time

According to eq. 1, jitter is directly related to the difference of the waiting times incurred by two consecutive packets. It is well-known that tail probabilities of the steady-state waiting time W in GI/GI/1 queues often have approximately an exponential form, i.e.

$$P\,[W > x] = \alpha\, e^{-\eta\, x} \tag{2}$$

for suitably large x, where the decay rate η and the constant α are fixed positive real numbers. Such approximations are largely available in the queueing litterature [2], [22], [31], [29] and are known to often have a remarkable quality even for small values of x.

In the sequel, we will use the approximation proposed by Abate et al. [2]. Let V be a generic service time and U/ρ be a generic interarrival time in the model with traffic intensity ρ. Thus, U and V are both fixed random variables with mean 1. Let u_k and v_k denote the kth moment of U and V, respectively. Thus, $u_1 = v_1 = 1$, and the square coefficient of variations (SCV) are $c_a^2 = u_2 - 1$ for the arrival process and $c_s^2 = v_2 - 1$ for the service time process. The approximation is as follows,

$$\eta = \frac{2\,(1-\rho)}{c_a^2 + c_s^2}\ [1 - (1-\rho)\,\eta^*]$$

where,

$$\eta^* = \frac{[2v_3 - 3c_s^2(c_s^2+2)] - [2u_3 - 3c_a^2(c_a^2+2)]}{3\,(c_a^2+c_s^2)^2}$$

The constant α is an approximation of $\eta\, E[W]$ given by,

$$\alpha = \frac{\eta\,\rho}{1-\rho}\left[\frac{c_a^2+c_s^2}{2} - (1-\rho)\,g(c_a^2, c_s^2, u_3)\right]$$

where,

$$g(c_a^2, c_s^2, u_3) = \left(\frac{2u_3/3 + (c_a^2+1)(c_s^2-1)}{2\,(c_a^2+c_s^2)} \ - \ c_a^2\right)$$

To assess the quality of this approximation, let us consider the simple case of a M3/D3/1 queue, i.e. $a(t) = \lambda\, e^{-\lambda t}$ and $b(x) = p_1\, u_0(x - T_1) + p_2\, u_0(x - T_2) + p_3\, u_0(x - T_3)$, where u_0 denotes the unit impulse. The following values are assumed: $T_1 = 1$, $p_1 = 0.6$, $T_2 = 5$, $p_2 = 0.3$, $T_3 = 10$ and $p_3 = 0.1$. For these values, the first three moments of the service time distribution are 3.1, 18.1 and 138.1, and the SCV is 0.88345. Figures 1 and 2 show the quality of approximation 2 for $\rho = 0.4$ and $\rho = 0.8$.

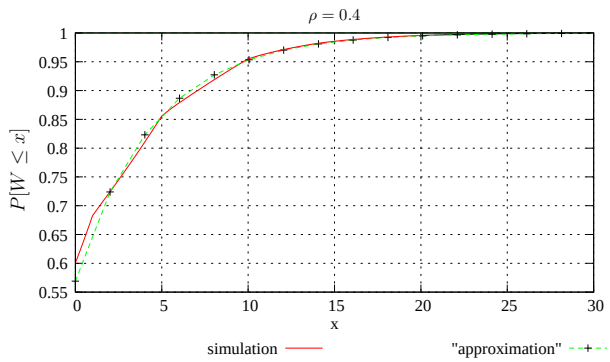

Fig. 1. Exponential approximation of $P(W \leq x)$ for $\rho = 0.4$.

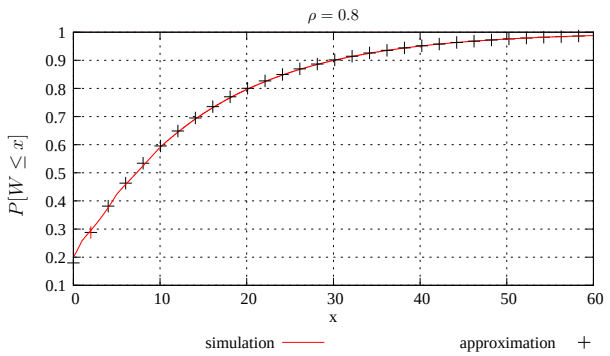

Fig. 2. Exponential approximation of $P(W \leq x)$ for $\rho = 0.8$.

IV. Single node case

In this section, we consider a single queue and, to simplify notations, we drop the index k of the queue. If the queue is not heavily loaded, the packet C_1 enters the queue when C_0 has already departed, i.e. $W_0 \leq T'$ where

$T' = T - D/\mu$. In this case, W_1 and W_0 can be considered as two independant and identicaly distributed random variables whose distribution is the distribution of the steady-state waiting time of the queue. According to approximation 2, the pdf of the waiting time is,

$$f_W(x) = (1-\alpha)\, u_0(x) \;+\; \alpha\, \eta\, e^{-\eta\, x} \tag{3}$$

where η and α are given by the formulae in section III using the first three moments of the arrival process and the first three moments of the service time process of the background traffic. By conditioning on the event $W_0 = y \leq T'$, we have the following approximation:

$$P\,[W_1 > x + y] = \alpha\, e^{-\eta\,(x+y)} \tag{4}$$

Using this approximation, it can be shown that the cumulative distribution function (CDF) of $\Delta = W_1 - W_0$ is given by,

$$F_\Delta(x) \approx \begin{cases} 0 & , x \leq -T' \\ 1 - h(T')\,[1 - q_1(x, T')] & , -T' \leq x \leq 0 \\ 1 - h(T')\, q_2(x, T') & , x \geq 0 \end{cases} \tag{5}$$

where,

$$\begin{aligned} h(T') &= \frac{1}{1 - \alpha e^{-\eta T'}} \\ q_1(x, T') &= \alpha\, e^{\eta x} + \frac{\alpha^2}{2}\,(e^{\eta x} - e^{-\eta(2T'+x)}) \\ q_2(x, T') &= \alpha\, e^{-\eta x}\left[1 - \frac{\alpha}{2}\,(1 + e^{-2\eta T'})\right] \end{aligned}$$

To assess the quality of this approximation, let us consider again the M3/D3/1 queue described in section III. The service rate is 1 data unit per second. Assume a periodic packet stream, with constant interarrival time $T = 50$ and constant packet size $D = 1$, is multiplexed with the packets of the Poisson background traffic. Figures 3 and 4 compare the CDF obtained using eq. 5 and the empirical CDF obtained by event-driven simulations for several values of $\rho = 3.1 \times \lambda^{\text{bg}} + 1/T$. It can be seen that the approximation is accurate for $\rho \leq 0.6$ and becomes rather crude for ρ=0.8 (the independance assumption does not hold in this case).

Let $\phi = 2\,\alpha\,(1 - \alpha/2)$. According to 5, the pdf of Δ can be approximated as follows,

$$f_\Delta(x) \approx h(T') \begin{cases} 0 & , x \leq -T' \\ \frac{\phi\eta}{2}\, e^{\eta x} - \frac{\alpha^2\eta}{2}\, e^{-\eta(2T'+x)} & , -T' \leq x < 0 \\ 1 \;-\; \phi & , x = 0 \\ \left[\frac{\phi\eta}{2} \;-\; \frac{\alpha^2\eta}{2}\, e^{-2\eta T'}\right] e^{-\eta x} & , x > 0 \end{cases} \tag{6}$$

By definition, the jitter of the tagged packet stream is,

$$\begin{aligned} J &= E\,[\,|\,W_1 - W_0\,|\,] \\ &= -\int_{-T'}^{0} x\, f_\Delta(x)\, dx \;+\; \int_0^\infty x\, f_\Delta(x)\, dx \end{aligned}$$

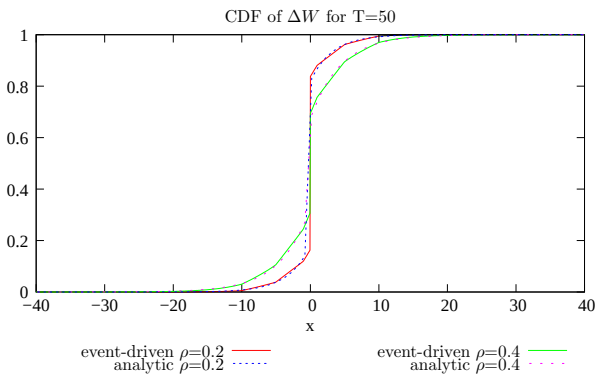

Fig. 3. Approximation of the CDF: ρ=0.2 and ρ=0.4.

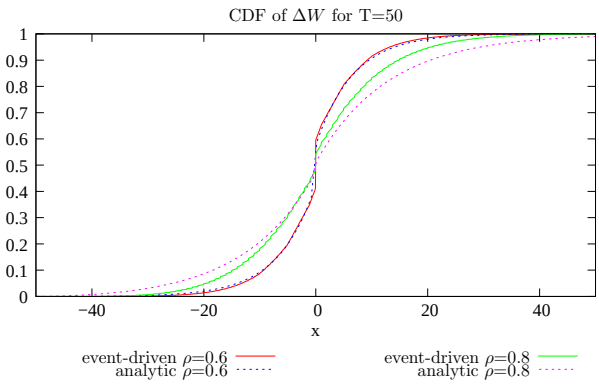

Fig. 4. Approximation of the CDF: ρ=0.6 and ρ=0.8.

After some algebra, it yields,

$$J = \frac{\alpha}{\eta}\; \frac{(2-\alpha) \;-\; e^{-\eta T'}\left[\eta T' + 1 - \alpha + \alpha\, e^{-\eta T'}\right]}{1 - \alpha\, e^{-\eta T'}} \tag{7}$$

where $T' = T - D/\mu$. Note that α/η is the average waiting time that would be incurred by a packet arriving at random if there was only packets of the background traffic.

A. *Results*

A.1 Example 1

Let us consider again the M3/D3/1 queue described in section III. Assume a periodic packet stream, with constant interarrival time T and constant packet size $D = 1$ data unit, is multiplexed with the packets of the Poisson background traffic. In the figures below, we present comparisons of the jitter given by formula 7 and the empirical jitter given by event-driven simulations. We assume that $\mu = 1$ data unit/s and that the utilization rate of the queue is kept constant by increasing $\lambda_T = 1/T$ and decreasing accordingly the intensity $\lambda^{bg} = \rho - \lambda_T$ of the background traffic.

Figure 5-8 show that the jitter approximation is very accurate for light or moderate utilization rates. The approximation is rather crude when the queue is saturated, as illustrated by the case $\rho = 0.8$.

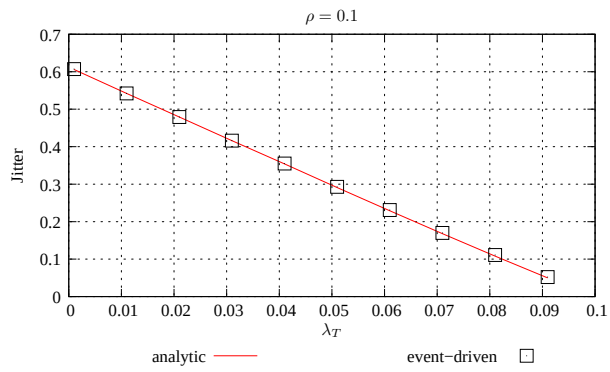

Fig. 5. Single M3+D/D3+D/1 queue: $\rho = 0.1$.

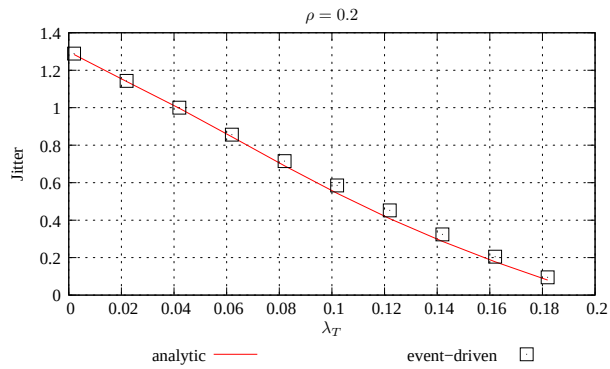

Fig. 6. Single M3+D/D3+D/1 queue: $\rho = 0.2$.

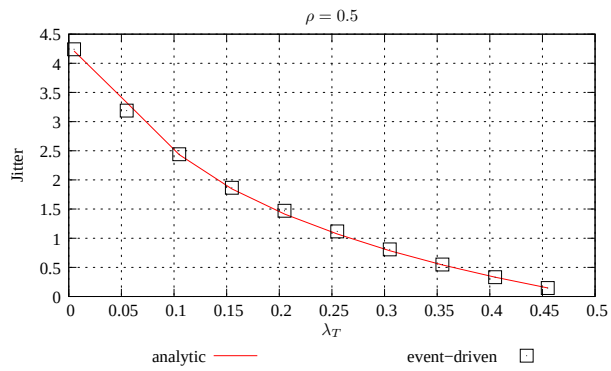

Fig. 7. Single M3+D/D3+D/1 queue: $\rho = 0.5$.

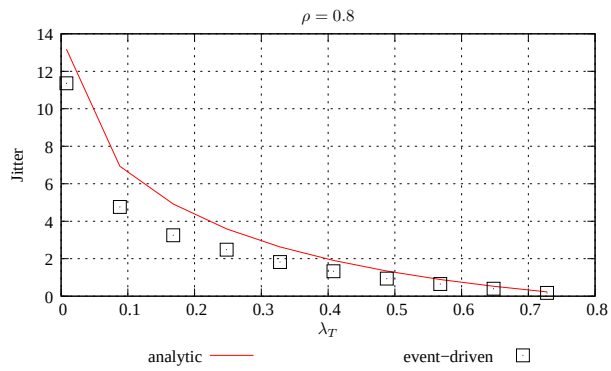

Fig. 8. Single M3+D/D3+D/1 queue: $\rho = 0.8$.

Codec	Packet Size (Bytes)	T_{ON} (s)	T_{OFF} (s)	R (packets/s)
G.711	136	0.352	0.650	1/0.012
G.726	104	0.352	0.650	1/0.016
G.729	70	0.352	0.650	1/0.030

TABLE I

PARAMETERS OF VOICE CODECS.

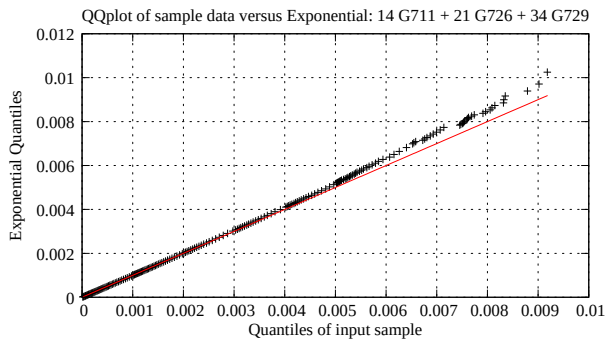

Fig. 9. QQ plot for a mixture of 14 G711, 21 G726 et 34 G729.

A.2 Example 2

We assume that the background traffic is the aggregation of several digitized voice traffics, generated using two different standard speech coders : G.726 and G.729. The standard practice is to model voice packet streams as an ON-OFF processes [4], [17], [30], with parameters depending on the codec. Let T_{ON} be the mean speech activity period, T_{OFF} be the mean idle period and R be the constant packet rate during an On-period. The values of these parameters are given for each voice codec in table I.

For each one of these voice codecs, the probability of the ON state is $P_{ON} = T_{ON}/(T_{ON} + T_{OFF}) = 0.3513$. Let n_{G726} and n_{G729} be the number of G.726 flows and G.729 flows, respectively. The total packet rate of the background traffic is therefore,

$$\lambda^{\text{bg}} = P_{ON}\ [n_{G726}\, R_{G726}\ +\ n_{G729}\, R_{G729}]$$

In the following, we consider two different mixtures of voice traffics. The first mixture assumes 10% of G.726 voice flows and 90% of G.729 voice flows. The second mixture assumes 50% of each type.

We assume that the background traffic can be modeled as a Poisson packet stream. This assumption is a rather crude approximation when the number of voice flows is low, but it becomes more and more accurate as the number of voice flows is increased. This is illustrated by the QQ plot in figure 9, where the sample data are the interarrivals of a superposition of voice traffics. Our experiments show that when the number of voice traffics is greater than 15 or 20 for each voice codec, the Poisson assumption holds.

The packet size distribution is a two-point distribution. The parameters of this distribution are given by the packet size of each codec and by the probability for an incoming

packet to be of type G.726 or G.729:

$$p_{G726} = \frac{n_{G726} R_{G726} P_{ON}}{\lambda_{bg}} \quad p_{G729} = \frac{n_{G729} R_{G729} P_{ON}}{\lambda_{bg}} \quad (8)$$

It is therefore easy to compute the first three moments of the packet-size distribution and thus the parameters α and η.

The packets of this aggregation of voice traffics are multiplexed with a single G.711 voice traffic on a link. The capacity of the link is $\mu = 1920$ kbps $= 245760$ Bytes/s.

For several packet-rates of the background traffic, the jitter of the single G.711 voice traffic has been computed using approximation 7 (J_{ana}), and compared to the empirical jitter (J_{eve}) obtained by event-driven simulations. Results are synthetized in figure 10. For the first (resp. second) mixture, the relative error is below 1.7% (resp. 2.8%) if the utilization rate is less than 54%. It is 6.5% (resp. 13.5%) if the utilization rate is 73.6%.

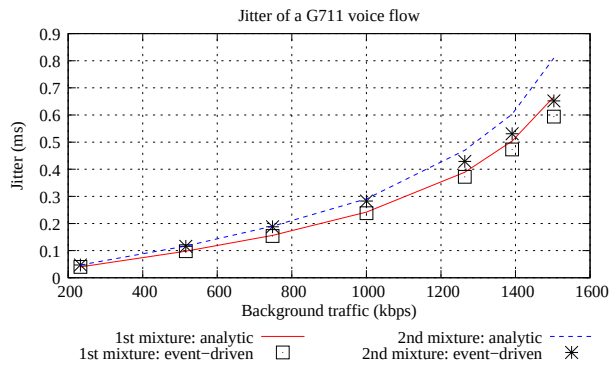

Fig. 10. Jitter of a G.711 voice flow multiplexed with a mixture of voice traffics .

A.3 Example 3

Let us assume that the interarrival times of the background traffic have a two-phase hyperexponential distribution H2 with balanced means, i.e.

$$a^{\text{bg}}(t) = p_1 e^{-\lambda_1 t} + p_2 e^{-\lambda_2 t}, \; t \geq 0 \quad \text{with} \quad \frac{p_1}{\lambda_1} = \frac{p_2}{\lambda_2}$$

As for the SCV, we use c_a^2=2.0.The packet sizes of the background traffic have a Gamma distribution with mean $k\theta = 1$ and shape parameters $k = 1/2$, i.e.

$$b^{\text{bg}}(x) = x^{k-1} \frac{e^{-x/\theta}}{\theta^k \, \Gamma(k)} = (2\,\pi\,x)^{-1/2} \, e^{-x/2}, \quad x \geq 0$$

The first three moments of the packet size distribution are 1.0, 3.0 and 15.0, and its SCV is c_s^2=2.0. We assume that the service rate μ is 1 data unit/s.

The tagged traffic has a rate $\lambda_T = 1/T$ and a constant packet size $D = 1$ data unit. Figure 11 compares our jitter approximation 7 with the empirical jitter given by event-driven simulations for several values of the global utilization rate ρ. The approximation is less accurate than in the M/Dk/1 case, but still remains satisfactory for light or moderate utilization rates.

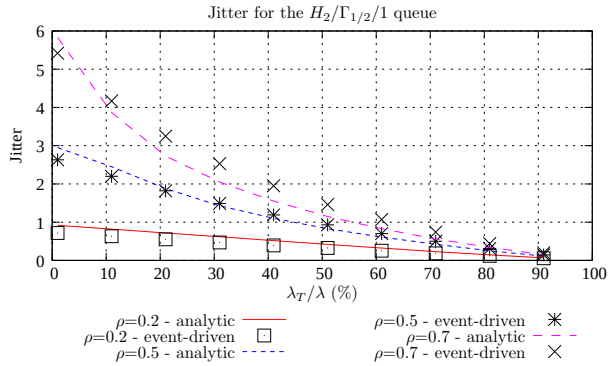

Fig. 11. Jitter approximation for the $H_2 + D/\Gamma_{1/2} + D/1$ queue.

It can be argued that the $\Gamma_{1/2}$ packet size density is reasonably close to an exponential packet size density. To make a more challenging comparison, we change the packet size distribution to a two-point distribution D2 and keep the H2 arrival process unchanged. The new packet size density is,

$$b^{\text{bg}}(x) = p_1 \, u_0(x - D_1) \; + \; p_2 \, u_0(x - D_2) \quad x \geq 0$$

with $p_1 = 0.019607843$, $p_2 = 0.980392157$, $D_1 = 11.0$ and $D_2 = 0.8$. The SCV of the packet size is still 2.0, and the first three moments are 1.0, 3.0 and 26.6. Figure 12 provides the same comparison than figure 11, but for the new packet size density. The same conclusion holds.

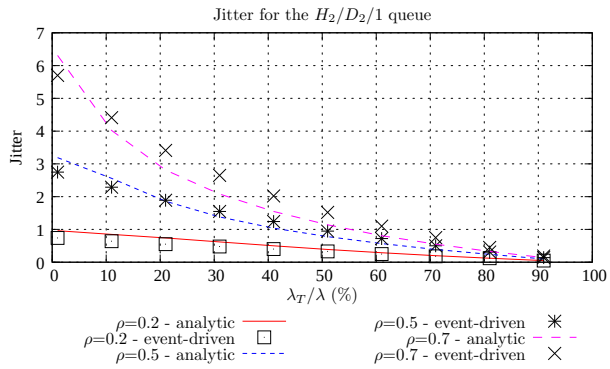

Fig. 12. Jitter approximation for an $H_2 + D/D2 + D/1$ queue.

B. *Approximation of the probability density function*

Recall that $T' = T - D/\mu$. In the sequel, we make the assumption that the terms weighted by $e^{-\eta T'}$ can be neglected. This assumption will hold if the offered load of the tagged traffic is low with respect to the service rate and if the queue is not saturated. If it holds, eq. 6 becomes,

$$f_\Delta(x) \approx \frac{\phi\eta}{2} e^{-\eta|x|} \; [1 - u_0(x)] \; + \; (1 - \phi) u_0(x) \; , \; x > -T' \quad (9)$$

and $f_\Delta(x) = 0$ if $x \leq -T'$. Figure 13 show that this approximation is very accurate for several values of the utilization rate. The unit impulse at $x = 0$ is not shown for clarity. We have assumed a Poisson arrival distribution and a three-point packet size distribution (the parameters of these distributions are given in section IV-A.1). The value of T is 50.0.

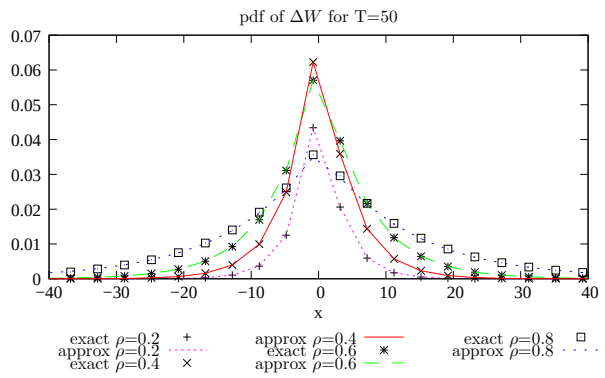

Fig. 13. Approximation of $f_\Delta(x)$.

This approximation will be used in the next section for the analysis of the multiple node case.

V. Multiple node case

In this section, we consider the multiple-node case. Knowing the arrival distribution and the packet size distribution of the background traffic for each node $k = 1, \ldots, n$, the parameters α_k, $\phi_k = 2\alpha_k(1-\alpha_k/2)$ and η_k can be computed for each node.

Let T_k, $k = 1, \ldots, n$ be the time between arrivals of C_0 and C_1 at node k. We obviously have,

$$T_1 = T \quad \text{and} \quad T_k = T + \sum_{j=1}^{k-1} \Delta_j \quad k = 2 \ldots n$$

Let us denote by $f_{\Delta_k|T}(x)$ the pdf of Δ_k conditioned on $T_k = T$ for $k = 1, \ldots, n$. According to formula 9, we have:

$$f_{\Delta_k\,|\,T}(x) = \begin{cases} 0 & , x \leq -T'_k \\ (1-\phi_k) & , x = 0 \\ \frac{\phi_k \eta_k}{2}\, e^{-\eta_k |x|} & , \text{otherwise} \end{cases} \tag{10}$$

where $T'_k = T - D/\mu_k$.

In the sequel, we assume $\eta_i \neq \eta_j$, $i, j = 1, \ldots, n$. The case $\eta_i = \eta_j$ can be handled by continuity. Let us define, for $j = 1, \ldots, n$, the coefficients K_j^n by:

$$K_j^n = \prod_{i=1, i\neq j}^{n} \left(1 - \phi_i \frac{\eta_j^2}{\eta_j^2 - \eta_i^2}\right)$$

According to eq. 1, the end-to-end jitter is given by,

$$J_{[1\ldots n]}(T) = E\left[\, |\sum_{k=1}^{n} \Delta_k\, | \,\right]$$

To compute the end-to-end jitter, we will derive an analytical approximation for the pdf $f_{\Delta_1+\ldots+\Delta_n}(x)$ of the sum $\Delta_1 + \ldots + \Delta_n$. The key observation is that the following recursion holds,

$$\begin{aligned} f_{\Delta_1+\ldots+\Delta_{n+1}}(x) \quad = \quad & \int_{-\infty}^{+\infty} f_{\Delta_1+\ldots+\Delta_n}(y) \times \\ & f_{\Delta_{n+1}|T+\Delta_1+\ldots+\Delta_n}(x-y|T+y)\, dy \end{aligned}$$

Using this recusion and approximation 10, the following proposition can be proved.

Proposition 1: The pdf $f_{\Delta_1+\ldots+\Delta_n}(x)$ is given by,

$$f_{\Delta_1+\ldots+\Delta_n}(x) \approx \begin{cases} 1 - \sum_{j=1}^{n} \phi_j\, K_j^n & , x = 0 \\ \sum_{j=1}^{n} K_j^n\, f_{\Delta_j\,|\,T}(x) & , x \neq 0 \end{cases}$$

Proposition 1 allows to derive a very simple expression of the end-to-end jitter, as stated below.

Proposition 2: The jitter introduced by nodes $1, \ldots, n$ can be approximated by,

$$J_{[1\ldots n]}(T) = \sum_{j=1}^{n} K_j^n\, J_j(T) \tag{11}$$

where $J_j(T)$ is the jitter that would be introduced by node j if the inter-arrival time at that node was T,

$$J_j(T) = \frac{\phi_j}{\eta_j}\left[1 \;-\; \frac{1}{2}\,(1 + \eta_j T'_j)\, e^{-\eta_j\, T'_j}\right]$$

with $T'_j = T - D/\mu_j$.

Hence, the end-to-end jitter can be seen as the weighted sum of jitters that would be introduced by each individual node if the inter-arrival time was T at each one of these individual nodes.

A. Results

B. Example 1

Let us consider $n = 5$ queues in tandem. All queues have a unit service rate, i.e. $\mu_k = 1$ data unit/s, $k = 1 \ldots n$. We assume Poisson arrivals and a three-point packet-size distribution for the background traffic of each queue. The parameters of these distributions are identical to those of section IV-A.1. The constant packet-size of the tagged traffic is $D = 1$ data unit.

Queues 2, 3, 4 and 5 have the same utilization rate ρ. The utilization rate of queue 1 is given by $\rho_1 = (1+\omega)\,\rho$, where ω can take the following values: -50%, -25%, 0%, 25% and 50%. Thus the first queue is either less loaded, equally loaded or more loaded than the other queues. Let ρ_{min} be the lowest utilization rate. Of course, ρ_{min} is either ρ_1 if $\omega \leq 0$ or ρ if $\omega \geq 0$.

For a given value of ω, we consider the following values of ρ: 0.2, 0.4 and 0.6. For each one of these values, λ_T is increased from 1% to 100% of ρ_{min} and the intensity λ_k^{bg} of the background traffic is decreased accordingly in order to keep constant utilization rates.

Figures 14 - 18 compare the jitter given by formula 11 with the jitter obtained by event-driven simulations for several values of ω and ρ. In most cases the approximation is fairly accurate, in particular for large values of T. The case $\rho = 0.6$ and ω=50% is not shown because in this case the approximation is far less accurate (37% error if λ_T is 1% of ρ_{min}) due to the violation of the independance assumption ($\rho_1 = 0.9$).

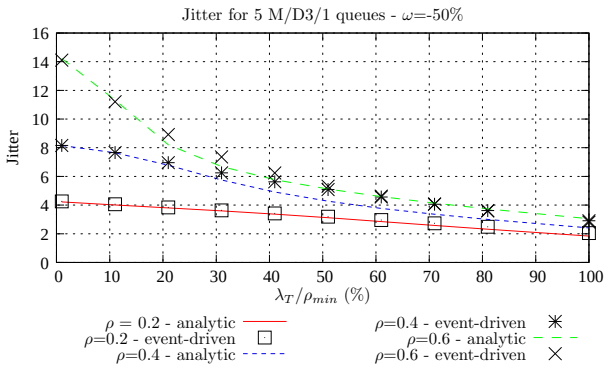

Fig. 14. Jitter approximation for ω = -50%.

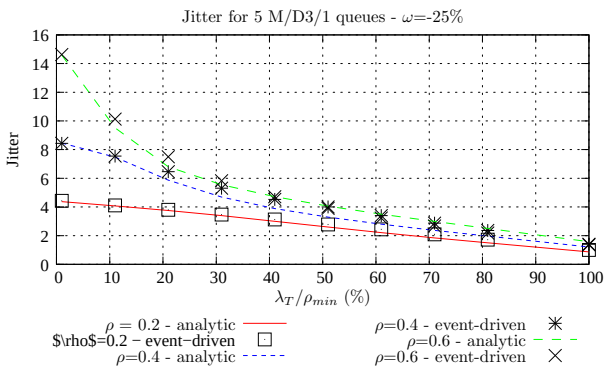

Fig. 15. Jitter approximation for ω = -25%.

C. Example 2

Let us consider $n = 5$ nodes in tandem and a G.711 voice flow sent from node 1 towards node 5. The first and last links have the same bandwidth, $\mu_1 = \mu_4$=1920 kbps. The bandwidth of the two other links is $\mu_2 = \mu_3$=7680 kbps.

We assume that the background traffic on the first and last link is the aggregation of 10% of G.726 voice flows and 90% of G.729 voice flows. The traffic on the two other links is the aggregation of 50% of each type. Moreover, since links 2 and 3 have a much higher bandwidth, it is assumed that the rate of their background traffic is approximately

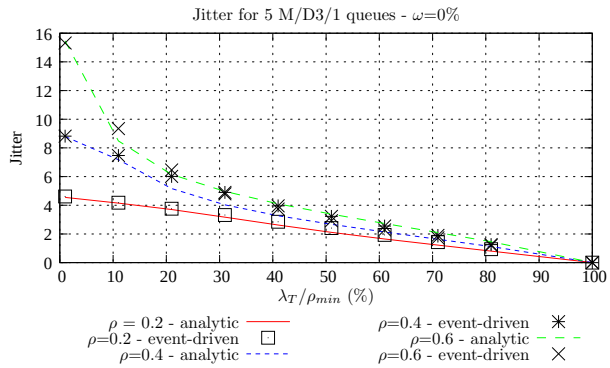

Fig. 16. Jitter approximation for ω = 0%.

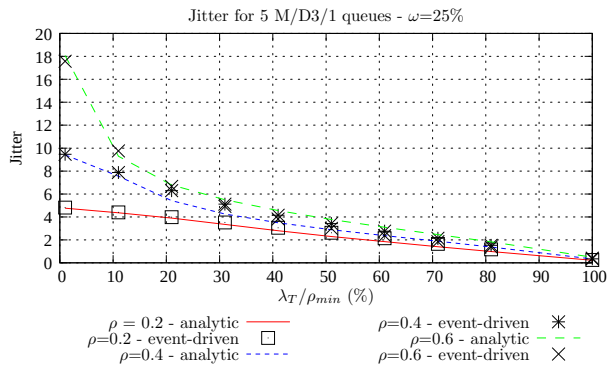

Fig. 17. Jitter approximation for ω = 25%.

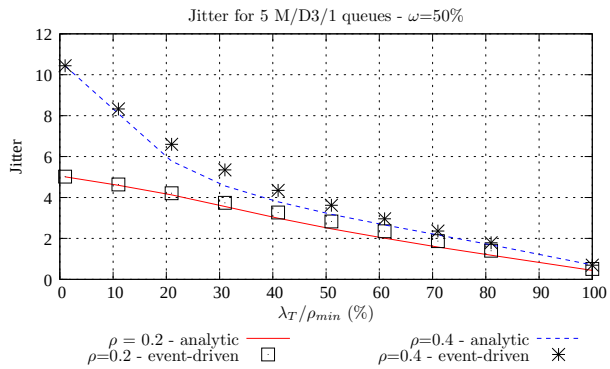

Fig. 18. Jitter approximation for ω = 50%.

3 times the rate of the background traffic on links 1 and 4 (integer values of the number of G.726 and G.729 voice flows are required to perform the event-driven simulation). As in section IV-A.2, the analytic model approximates the aggregation of voice flows by a Poisson arrival process and a two-point service time distribution.

Figure 19 plots both the jitter given by the analytic model and the empirical jitter given by event-driven simulations according to the rate of the background traffic on the first (and last) link. It can be seen that the analytic approximation is fairly accurate. The error is below 1% when the utilization rate of the first queue is less than 46.2%. It is only 2.08% if this utilization rate is 53.4%, and 3.46% if the utilization rate is 59.9%.

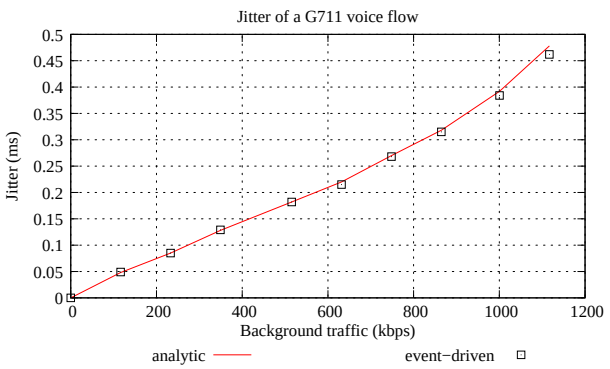

Fig. 19. Jitter approximation for a G.711 voice flow passing through 4 queues.

D. Example 3

Let us consider $n = 5$ queues in tandem. All queues have a unit service rate, i.e. $\mu_k = 1$ data unit/s, $k = 1 \ldots n$. We assume hyperexponential arrival distributions and Gamma packet-size distribution for the background traffics as in section IV-A.3. The constant packet-size D of the tagged traffic is 1 data unit.

All queues have the same utilization rate $\rho = 0.4$. Figure 20 compares the jitter obtained with formula 11 and the jitter obtained by event-driven simulations. Even if less accurate than in the Poisson case, the approximation is still satisfactory.

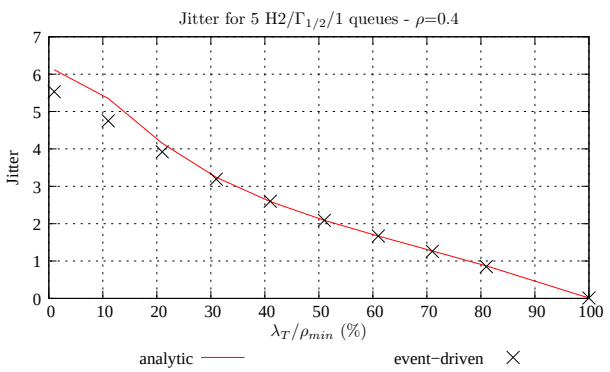

Fig. 20. Jitter approximation for 5 $H_2/\Gamma_{1/2}/1$ queues in tandem.

As in section IV-A.3, we change the gamma packet-size distribution for a more challenging two-point packet size distribution. Figure 21 shows that the approximation is still satisfactory.

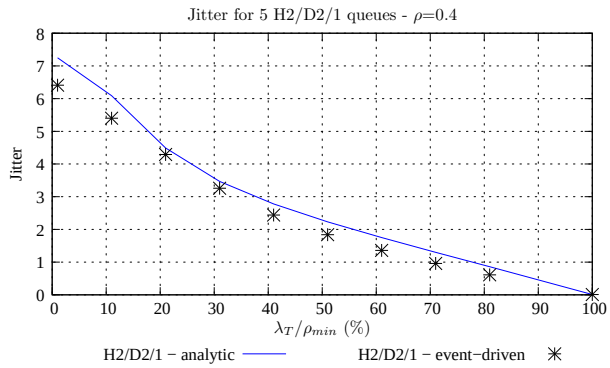

Fig. 21. Jitter approximation for 5 $H_2/D2/1$ queues in tandem.

VI. Conclusion

Real-time CBR traffics, e.g. traffics generated by most voice codecs and by some video codecs, are expected to be a significant fraction of the total traffic in the next generation Internet. A fundamental measure of performance for those traffics is the jitter, i.e. the distorsion of the original periodic pattern of the traffic due to the multiplexing operations performed by routers.

In this paper, we have derived an analytic approximation for the evaluation of end-to-end jitter incurred by a periodic packet stream with a constant packet size. Comparisons with event-driven simulations have shown that the proposed approximation is fairly accurate both in the single-node and in the multiple-node cases.

With respect to the numerical methods proposed by prior works, our approximation has the advantage of being very simple thus providing insight on jitter in IP networks. Moreover, such simple formulae are really required if one considers the simulation of large networks with thousands of real-time flows or the iterative optimization of networks under QoS constraints.

Another advantage of our approximation is that it is fairly general since the only assumption regarding the background traffic is that arrivals and service-times are governed by renewal processes. Moreover, this assumption can probably be removed since exponential approximations of the tail probabilities of the waiting time have also been derived for more general queueing models (e.g BMAP/GI/1 queues [1]).

The main limit of our approximation lies in the independance assumption regarding the waiting times of two consecutive packets. Our results show that this assumption holds in light to moderate traffic conditions, i.e. for utilization rates lower than 70% or 80%. Our future research will focus on extending this approximation to heavy traffic conditions.

References

[1] Abate J., Choudhury G.L., Whitt W., *Asymptotics for steady-state tail probabilities in structured Markov queueing models* . Stoch. Models 10, 99-143, 1994.

[2] Abate J., Choudhury G.L., Whitt W. *Exponential approximations for tail probabilities in queues, I: waiting times* . Operations Research, vol 43, No 5, 1995.

[3] Boyer, P. E., Guillemin, F. M., Servel, M. J., Coudreuse, J.-P. *Spacing Cells Protects and Enhances Utilization of ATM Network Links* , IEEE Network, vol. 6, 1992.

[4] Brady, P. T. *A Model for Generating ON-OFF Speech Patterns in Two-Way Conversations*, Bell System Technology Journal, Vol. 48, 1969.

[5] Alshaer, H., Horlait, E. *Expedited Forwarding End to End Delay Jitter in the Differentiated Services Networks*, Proceedings of HSNMC 2004, 2004.

[6] Cruz, R., *A calculus for network delay, Part I: network elements in isolation*, IEEE Trans. on Inf. Theory, vol. 37, 1991.

[7] Cruz, R., *A calculus for network delay, Part II: network analysis*, IEEE Trans. on Inf. Theory, vol. 37, 1991.

[8] Parekh, A., Gallager, R., *A generalized processor sharing approach to flow control in integrated services networks: the single node case*, IEEE/ACM Trans. on Networking, vol. 1, 1993.

[9] Parekh, A., Gallager, R., *A generalized processor sharing approach to flow control in integrated services networks: the multiple node case*, IEEE/ACM Trans. on Networking, vol. 1, 1993.

[10] Kurose, J., *On computing per-session performance bounds in high-speed multi-hop computer networks*, Proceedings ACM SIGMETRICS'92, 1992.

[11] Yaron, O., Sidi, M., *Calculating performance bounds in communication networks*, Proceedings IEEE INFOCOM'93, San Francisco, 1993.

[12] Cidon, I., Khamisy, A., Sidi, M., *Dispersed messages in discrete-time queues: delay, jitter and threshold crossing*, IEEE INFOCOM, Toronto, Canada, 1994.

[13] Ferrari, D., *Delay jitter control scheme for packet-switching internetworks*, Computer Communications, 15(6):367-373, 1992.

[14] Guillemin, F., Roberts, J. W., *Jitter and bandwidth enforcement*, IEEE GLOBECOM, pp 261-265 (9.3), Phoenix, Arizona, 1991.

[15] Guillemin, F., Monin, W., *Limitation of cell delay variation in ATM networks*, Proceedings of ICCT'92, Beijing, 1992.

[16] Guillemin, F., Boyer, P., Dupuis, A, Romoeuf, L., *Peak rate enforcement in ATM networks*, IEEE INFOCOM, Florence, Italy, 1992.

[17] Heffes, H., Lucantoni, D. M., *A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance*, IEEE Journ. of Selected Areas in Communications, pp. 856-867, Sept. 1986.

[18] Lam, S., Xie, G., *Burst scheduling: architecture and algorithm for switching packet video*, IEEE INFOCOM, Boston, Massachussets, 1995.

[19] Landry, R., Stavrakakis, I., *Traffic shaping of a tagged stream in an ATM network: approximate end-to-end analysis.* IEEE INFOCOM, Boston, Massachussets, 1995.

[20] Landry, R., Stavrakakis, I., *Study of delay jitter with and without peak rate enforcement.* IEEE/ACM Transactions on Networking, vol. 5:4, 1997.

[21] Li, S. Q., Fulton, C., *Delay jitter first-order statistics of general traffic on high-speed networks.* IEEE ICC, Dallas, Texas, 1996.

[22] Liu, Z., Nain, P., Towsley, D., *Exponential bounds with applications to call admission.* Journal of the ACM, Vol. 44, No 3, 1997.

[23] Matragi, W., Bisdikian, C., Sohraby, K., *Jitter calculus in ATM networks: single node case*, Proceedings IEEE Infocom, Toronto, Canada, 1994.

[24] Matragi, W., Sohraby, K., Bisdikian, C., *A Framework for Jitter Analysis in Cell Based Multiplexors*, Performance Evaluation, vol. 22, 1996.

[25] Matragi, W., Sohraby, K., Bisdikian, C., *Jitter calculus in ATM networks: multiple node case.* IEEE/ACM Transactions on Networking, vol. 5, 1997.

[26] Privalov, A., Sohraby, K., *Per-stream jitter analysis in CBR ATM multiplexors*, IEEE/ACM Transactions on Networking, vol. 6:2, 1998.

[27] Roberts, J. W., *Jitter due to an ATM multiplex - application to peak rate policing*, Technical Report COST 224-35, RACE Cost Project, 1989.

[28] Roberts, J. W., Guillemin, F., *Jitter in ATM networks and its impact on peak rate enforcement*, Performance Evaluation, 16:35-48, 1992.

[29] Seelen, L. P., Tijms, H. C., Van Hoorn, M. H., *Tables for Multi-Server Queues*, North-Holland, Amsterdam. 16:35-48, 1985.

[30] Sriram, K., Whitt, W., *Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data*, IEEE Journal of Selected Areas in Communications, pp. 833-846. 16:35-48, Sept. 1986.

[31] Tijms, H. C., *Stochastic Modeling and Analysis: A Computational Approach*, John Wiley, New York. 16:35-48, 1986.