



Statistical performance guarantees for streaming flows using expedited forwarding

Thomas Bonald, Alexandre Proutière, James Roberts

► To cite this version:

Thomas Bonald, Alexandre Proutière, James Roberts. Statistical performance guarantees for streaming flows using expedited forwarding. Infocom, 2001, Anchorage, United States. 10.1109/INF-COM.2001.916304 . hal-01244778

HAL Id: hal-01244778

<https://hal.science/hal-01244778>

Submitted on 16 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical Performance Guarantees for Streaming Flows using Expedited Forwarding

T. Bonald, A. Proutière, J. W. Roberts

France Telecom R&D

{thomas.bonald,alexandre.proutiere,james.roberts}@francetelecom.com

Abstract—We suggest that satisfactory statistical performance guarantees for streaming flows can be fulfilled when their packets receive expedited forwarding in non-preemptive priority queues. This relies on the conjecture that jitter remains negligible in the network such that performance measures can be bounded by assuming flows constitute Poisson arrival processes of MTU sized packets. We provide analytical and simulation evidence in support of this conjecture and show how it leads to simple engineering rules for both constant and variable rate streaming traffic.

I. INTRODUCTION

In this paper we consider the possibility of providing statistical performance guarantees suitable for audio and video streaming applications in the Internet. Such guarantees are formulated in terms of packet loss probabilities and packet delay bounds which are only exceeded with a certain small probability. They generally allow higher link utilization than deterministic guarantees and can also result in simpler traffic management procedures.

Streaming flows can have differing QoS requirements. In this paper, however, we assume that a single class of service is provided offering quality of service suitable for the most exigent application. A particularly demanding application is telephone quality voice over IP (VoIP) for which, it is argued in [6], packet queuing delay in the backbone should not exceed 10 ms.

We envisage that streaming flows are handled in network nodes as an aggregate using the expedited forwarding (EF) per-hop behaviour of the Diffserv model [10]. More precisely, we assume EF is realized using non-preemptive priority queuing with streaming packets having highest priority and being handled in a FIFO queue. We mainly consider the multiplexing of constant rate flows manifested by an initially periodic stream of constant size packets. However, the analysis carries over to variable rate flows on assuming admission control is employed to ensure that, with sufficiently high probability, the combined input rate is less than the service rate. The packet size of any flow is bounded by the maximum transfer unit, MTU. An essential assumption is that all flows are statistically independent at the network ingress.

Evaluating delays is complicated by the phenomenon of jitter. The exact periodicity of constant rate flows is lost as soon as these are multiplexed in a common queue and jitter accumulates as flows advance in the network. Worst case scenarios identified in the context of deterministic guarantees suggest that maximum delays can be unbounded, even when all flows are initially shaped to their nominal rate. To derive useful statistical performance bounds, it is important to understand how jitter impacts

traffic management functions like admission control, buffer sizing and traffic shaping.

It proves difficult to adequately quantify the jitter phenomenon. Parameters relating to the packet inter-arrival time distribution or to the difference in delay of successive packets, in particular, are of little practical use for the purpose of network engineering. The approach developed in the present paper is simply to characterize the jitter affecting a flow by specifying when we can ignore it for the purposes of traffic management. By ignoring jitter we mean that the same engineering rules apply independently of whether the flows are perfectly periodic, as at the ingress, or have suffered variable delay at prior multiplexing stages.

A worst case assumption for computing delay in a FIFO queue offered a superposition of independent periodic flows is to assume the latter constitute a Poisson process of packets of size MTU (see Section IV below). This allows the computation of performance bounds on loss and delay enabling buffer sizing and the definition of limiting load levels without the need to specify individual flow characteristics. We adopt this as the basis of traffic management rules and define jitter as being “negligible” if the same rules can be applied at any network node. In other words, a flow has “negligible jitter” if it is better than a Poisson stream of MTU-sized packets in the sense that performance in a stable FIFO queue would be worse if that flow were replaced by the Poisson/MTU stream. The usefulness of this notion depends on the following conjecture holding true:

The NJ conjecture: *If the network realizes EF using priority queuing, if flows eligible for EF have negligible jitter at the ingress with respect to the Poisson/MTU reference process and if at every multiplexing stage within the network the sum of input rates is less than the service rate, then the flows retain the negligible jitter property throughout the network.*

The NJ conjecture is essentially based on intuition. The main contribution of this paper is to present analytical and simulation results to support this intuition. We also show how assuming the NJ conjecture allows the definition of simple traffic management rules leading to acceptable performance guarantees for streaming applications.

In the next section we discuss related literature. We then contrast the traffic management rules necessary to ensure deterministic and statistical performance guarantees, respectively, and show how the latter lead to acceptable end-to-end performance even for the most demanding streaming applications. In Section IV we discuss the difficult question of comparing stochastic processes, introducing a number of orders and illustrating their application. Section V is devoted to a study of how jitter prop-

Partial funding for the research reported here was provided by the European Union framework programme ITEA through the project RTIPA.

agates in the network. Most of the analysis is concerned with a simple tandem network representing the path of a test flow but we also derive an asymptotic NJ result for general feed forward networks. Finally, in Section VI, we indicate how the negligible jitter conjecture can also be applied to variable rate flows.

II. RELATED WORK

A number of previous papers are particularly relevant to the present work. Our approach is largely inspired by the definition of negligible cell delay variation (CDV) in ATM networks and its analysis by Brichet *et al* [2] and Massoulié [11]. The concept of negligible CDV has been incorporated in ITU Recommendations on B-ISDN traffic engineering [9]. In the present paper we generalize the results of [2] and [11] to the case of variable packet size IP networks.

The study of CBR service in ATM networks by Grossglauser and Keshav [7] considers jitter through the phenomenon called “bunching”. The authors demonstrate by simulation that bunching has no significant impact on end-to-end delays. They conclude that FIFO service is adequate for CBR traffic and per-flow re-shaping at intermediate nodes is unnecessary. They derive a number of analytical heuristics for comparing processes based essentially on the second moment of the inter-arrival time and the index of dispersion for intervals (IDI). Our approach uses more general tools for comparing stochastic processes laying a sounder foundation for the analysis of “better than” orderings.

Sahni *et al* in [14] revisit the simulation experiments reported in [7] and generalize the traffic model to account for variable packet sizes. They reconfirm the adequacy of FIFO for constant size packets but suggest this may not be the case when packet sizes are very different from flow to flow. In particular, they observe that FIFO can result in considerably greater delay and delay variation for the smaller packets than per-flow weighted fair queuing. They take this observation as evidence that simple FIFO queuing may not be adequate in a network with widely different packet sizes. Our results and the NJ conjecture are not in contradiction with these results since we only claim packet delays are bounded by the delays that would be obtained if all packets had the maximum size MTU. We differ in the conclusion that this delay is prohibitive for the streaming flows of interest.

Additional simulation experiments are reported by Goyal *et al* [6]. These are tailored to the study of VoIP and incorporate background data traffic served either with lower priority or according to assigned weights in a class based fair queuing scheduler. They observe that priority queuing yields lower delays for the voice packets and that performance is considerably affected by the larger packets in the background traffic. In our study, bounds on end to end performance are calculated assuming worst case characteristics for the background traffic (MTU sized packets always present).

Finally, Guérin and Pla consider how jitter acquired in a network path affects the conformity of a flow with respect to its initial token bucket parameters [8]. They show that a limited amount of re-shaping is necessary to restore conformance at network boundaries. Our approach is quite different in that we do not consider conformance to one or several token buckets to be a useful characterization of jitter.

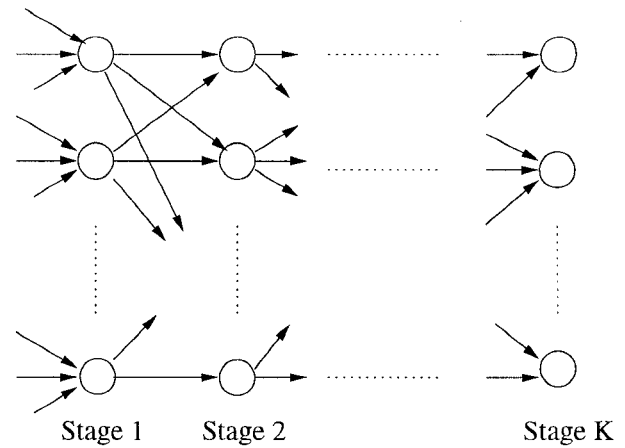


Fig. 1. Multistage network

III. STATISTICAL OR DETERMINISTIC PERFORMANCE BOUNDS

Guaranteed service in the Intserv model is based on deterministic performance bounds and generally relies on per-flow scheduling in network nodes. In this section we discuss the difficulty of providing similar deterministic performance guarantees in the Diffserv model where individual flows are not distinguished. We then contrast this approach with the traffic engineering rules possible if we can alternatively rely on the NJ conjecture to derive statistical performance bounds.

A. Deterministic bounds

Deterministic bounds on performance are derived by making worst case assumptions with respect to the simultaneity of arrivals from different flows and with respect to the accumulation of bursts of packets as the flows progress through the network. To illustrate how these worst case assumptions can be really pessimistic, we consider the simple network configuration depicted in Figure 1 consisting of a sequence of multiplexing stages realized by FIFO queues. Flows enter the network on the left and proceed through one queue in each stage before leaving at the right.

Assume N flows initially consisting of perfectly periodic streams of constant size packets are offered to the first stage queue. For a particular test flow, the worst case delay occurs when all flows emit a packet just ahead of its own packet and is equal to $N - 1$ packet times. Worst case jitter occurs when one packet is delayed like this while its successor suffers zero delay. This can occur if the flows having delayed the first packet suddenly stop sending before the second packet arrives. If the next packet is maximally delayed, the fourth minimally delayed and so on, the output process offered to a second stage queue is a periodic flow with packets arriving two by two.

A further worst case assumption for this network is that flows offered to a particular second stage queue have all been maximally jittered by the above process. Applying a similar worst case analysis accounting for simultaneous arrivals and alternating activity patterns shows that flows can acquire additional jitter at every stage. Worst case delay increases exponentially from stage to stage at a rate depending on how close the packets can

get (a function of N and the queue load).

In practice, even the first queue in the path of one flow will receive other flows which have already been jittered in previous multiplexing stages. The identification of worst case scenarios in this case becomes very complicated. Charny and Le Boudec [3], derive a general bound on end-to-end delay valid in a network where all flows conform on ingress to the same token bucket parameters and no flow follows a path including more than h queuing stages. This bound is derived more elegantly than in the above discussion but implicitly makes the same worst case assumptions. A consequence is that the delay bound increases rapidly with path length and link utilization. For a given path length, there is a limit utilization α^* beyond which the delay bound becomes infinite. The limit utilization is as low as $\alpha^* = 1/(h-1)$ if we do not take account of input rate limits. Clearly, allowed utilization becomes extremely small in large networks. It is also noteworthy that, even when finite, the delay bound increases linearly with the number of flows in the network.

Reliance on such deterministic bounds is not only unduly pessimistic with respect to possible link utilization but also implies the application of highly complex admission control and traffic management rules (including being aware of the number and rates of flows in multiple network paths). We note also that these rules are not even available for general network topologies and heterogeneous flow traffic parameters. It thus appears highly desirable to develop an alternative statistically based traffic management, as discussed below.

B. Statistical bounds

In this section we assume the NJ conjecture is valid and discuss the traffic management procedures which follow. If all flows have negligible jitter, the priority traffic at every node can be assimilated to a Poisson arrival process of MTU-sized packets. To derive worst case bounds we assume that the intensity of non-priority traffic is sufficiently high that whenever no high priority packet is present, the link is immediately occupied for the transmission of an MTU-sized lower priority packet.

B.1 A single multiplexing stage

Analysis of this queuing system yields the following expression for the Laplace transform of the distribution of queue size Q :

$$E[\exp -sQ] = \frac{1-\rho}{\rho} \frac{1 - e^{-MTU \cdot s}}{e^{-MTU \cdot s} + s/\lambda - 1}.$$

Rather than proceeding to invert this expression we apply the technique described in [13], page 392 to derive the asymptotic estimate:

$$\Pr[Q > x] \approx k \exp\{-rx/MTU\}$$

where r satisfies: $\rho(e^r - 1) - r = 0$ and $k = (1-\rho)/(\rho^2 e^r - \rho)$. Note that this approximation in fact provides an upper bound on the true waiting time distribution:

$$\Pr[Q > x] \leq \begin{cases} 1, & \text{if } x < x_{min} \\ k \exp\{-rx/MTU\}, & \text{if } x \geq x_{min}. \end{cases}$$

where $x_{min} = -MTU/r \log 1/k$. This yields accurate estimates of the tail probabilities of interest for buffer sizing. Table I

load	0.4	0.4	0.9	0.9
Pr[saturation]	10^{-3}	10^{-6}	10^{-3}	10^{-6}
buffer size	6	10	35	68

TABLE I
REQUIRED BUFFER SIZE IN PACKETS

hops	5	5	10	10
load	0.4	0.9	0.4	0.9
99% quantile	1.0	4.8	1.9	8.0
99.9% quantile	1.2	6.0	2.0	9.4

TABLE II
QUANTILES OF END-TO-END DELAY (MS)

shows required buffer size in numbers of MTU packets for a number of loads and target saturation probabilities.

B.2 End-to-end delay

The approximation can also be used to derive estimates of the end-to-end delay in a network of h stages. We assume all stages are identical, independent and of capacity C . The end-to-end delay is thus bounded by the sum of hx_{min}/C and a random variable having an Erlang- h distribution of mean $MTUh/r$. Quantiles of the end-to-end delay distribution are given in milliseconds in Table II assuming OC3 (155 Mbit/s) links and an MTU of 1500 bytes. Clearly the delays are very small for high speed links, even for long paths and high utilization.

The independence assumption is conservative since correlation due to dependence tends to reduce rather than increase end-to-end delays. We have verified this by simulation, see Figure 2.

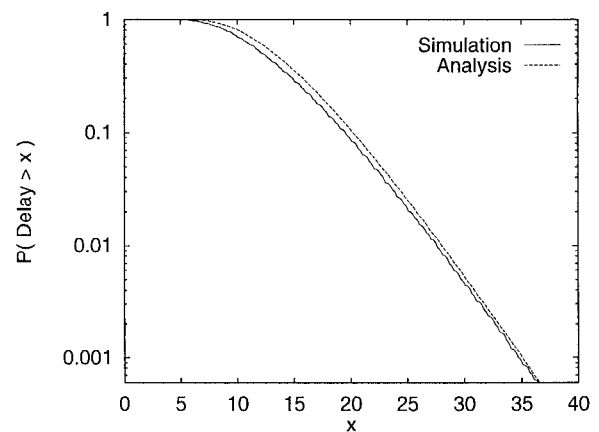


Fig. 2. Simulation: end-to-end delay of a Poisson through traffic contributing a load 0.1 passing through 4 queues in series with Poisson cross traffic at each queue. Total load 0.8. Analysis: end-to-end delay assuming independence.

B.3 Shaping and reshaping

The NJ conjecture states that flows retain the negligible jitter property if all flows have it at the network ingress. To ensure the latter condition is fulfilled it may be necessary to shape

the individual flows to their nominal constant rate, as recommended in the EF RFC [10]. According to the conjecture, any subsequent shaping beyond the ingress is superfluous if the network provider is sure that multiplexing is performed in the necessary manner. Shaping may, however, be necessary at an inter-network interface where the receiving provider does not trust the sending network to only forward flows with negligible jitter.

The shaping of an aggregate flow does not appear to serve any useful purpose. According to the conjecture, if the shaping rate is greater than the sum of the nominal flow rates and the latter have negligible jitter, the aggregate shaping does not introduce non-negligible jitter. However, it does introduce unnecessary delay (by emulating an additional heavily loaded link). It would also tend to exacerbate the jitter present in any component flow and cause any flow with non negligible jitter to “infect” the other flows included in the same aggregate. It is indeed the jitter in the component flows which impacts the performance of subsequent network queues. If reshaping is required it must be performed on flows or aggregates of flows which remain intact from ingress to egress.

C. Accounting for variable rate flows

A further advantage of applying statistical rather than deterministic performance guarantees is that the former allows the possibility of performing statistical multiplexing for variable rate flows. Deterministic guarantees for variable rate flows either require complex per-flow scheduling, as envisaged in the Intserv model, or a rate allocation equal to the peak rate if aggregate scheduling is imposed. Statistical delay bounds, on the other hand, can be guaranteed by applying a simple extension to the NJ conjecture, as discussed in Section VI. It is simply necessary to perform admission control to ensure a sufficiently small probability of the combined input rate exceeding the service rate.

IV. COMPARING STATISTICAL DELAYS

The NJ conjecture relies on the comparison of stochastic processes. In this section we review a number of results on the comparison of stationary marked point processes where the points are arrival epochs and the marks correspond to packet size. For the sake of simplicity, we do not take into account the effect of non-priority traffic when comparing processes through their induced queuing performance. We assume that if one process is better than another with respect to a single class queue, it will also lead to better performance in the priority queue.

A. Stochastic orders

Let $A = (N, \Sigma)$ denote a stationary marked point process where N is the point process of packet arrival epochs and Σ is the packet size. Let $V(A, c)$ be the virtual waiting time in a queue served at rate c and offered input process A . $A(0, t]$ denotes the total amount of work arriving in an interval of length t starting at the arbitrary point 0. In the following we introduce a number of stochastic orders, drawn from the literature or specifically adapted to present purposes.

Since the objective is to compare processes with respect to their performance in a queue, it is natural to define an order as follows.

Definition 1: Virtual waiting time ordering. We say $A \leq_{vwt} A'$ if and only if $\Pr[V(A, c) > x] \leq \Pr[V(A', c) > x]$, for $x > 0$ and for any c such that the queues are stable.

Unfortunately, it generally proves very difficult to apply this ordering in practice in view of the difficulty of calculating the virtual waiting time distribution.

Definition 2: Virtual waiting time convex ordering. $A \leq_{wcx} A'$ if and only if for all increasing convex functions h we have, $E[h(V(A, c))] \leq E[h(V(A', c))]$, i.e., $V(A, c) \leq_{icx} V(A', c)$, for any c such that the queues are stable.

It follows, in particular, that $A \leq_{wcx} A'$ implies all the moments of $V(A, c)$ are less than the respective moments of $V(A', c)$.

Definition 3: Strong variability ordering. $A \leq_{svar} A'$ if and only if for all convex functions h , $E[h(A(0, t))] \leq E[h(A'(0, t))]$, i.e., $A(0, t) \leq_{cx} A'(0, t)$, for $t > 0$.

For the particular case of renewal processes with constant marks (i.e., constant packet size), the strong variability ordering is satisfied if $\tau_A \leq_{cx} \tau_{A'}$ where τ denotes an inter-arrival time [1], [2]. If $A_i \leq_{svar} A'_i$ for $i = 1, 2$ then $(A_1 + A_2) \leq_{svar} (A'_1 + A'_2)$ (cf. Proposition 3 in [11]). Finally, by a classical result from Large Deviations theory [11], [4], the strong variability ordering implies the following weaker ordering based on asymptotic queuing performance.

Definition 4: Effective bandwidth ordering. For random variables for which the following limit exists, let $\alpha_A(c) = -\lim_{x \rightarrow \infty} \frac{1}{x} \ln P(V(A, c) > x)$, and define $A \leq_{EB} A'$ if and only if $\alpha_A(c) \geq \alpha_{A'}(c)$.

The inverse function α_A^{-1} is the effective bandwidth function giving the rate $c = \alpha_A^{-1}(\delta)$ necessary to achieve an asymptotic slope δ . $A \leq_{EB} A'$ implies A has smaller effective bandwidth than A' .

B. “Better than Poisson”

We first assume that packets have constant size so that we simply compare point processes of equal intensity. This is the context of the studies on cell delay variation in ATM networks reported in [2], [11].

Consider the superposition of periodic streams. In this case the performance measure in question is the virtual waiting time distribution when a server is offered a superposition of periodic streams of constant size packets where the phase of any stream is chosen at random. When all streams have the same rate, the considered system is the $N^*D/D/1$ queue for which an exact expression is available for the virtual waiting time distribution (e.g., [13]). Evaluations of this expression reveal that $\Pr[V > x]$ is an increasing function of N and tends in the limit to the corresponding value of the virtual waiting time distribution for the $M/D/1$ queue (see Figure 3). In other words, the N^*D/D process is better than Poisson in the $\leq_{vwt}$ ordering. When each stream has its own period, the corresponding system is the $\sum D_i/D/1$ queue for which there is an accurate approximation for the virtual waiting time distribution [13]. Numerical evaluations reveal that the $M/D/1$ queue again provides an upper bound on performance indicating that any superposition of independent periodic streams is better than Poisson.

In fact, it can be demonstrated that a periodic process of constant size packets is better than Poisson in the strong vari-

ability ordering (clearly, by Jensen's inequality, $h(1/\lambda) \leq \int_0^\infty h(x)\lambda e^{-\lambda x} dx$ for convex functions h showing that the deterministic renewal process is better than Poisson). The results following Definition 3 allow us to conclude that this order extends to superpositions of periodic flows. In this case, the numerical results of Figure 3 indicate that the strong variability order implies the sought-for virtual waiting time order.

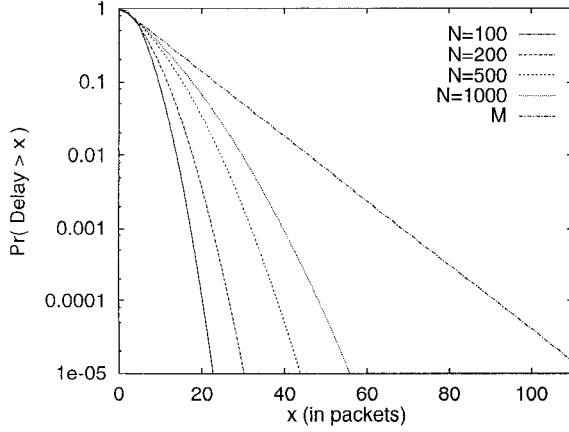


Fig. 3. Virtual waiting time distribution of the $N^*D/D/1$ queue. Load $\rho = 0.95$.

C. "Better than Poisson/MTU"

In the Internet, packets are of variable size bounded by the maximum transfer unit MTU and, to compare the performance of different types of flow, it is necessary to use results for marked point processes.

The $\sum D_i^{X_i}/D/1$ queue (a superposition of independent streams of periodic arrivals of size X_i) models constant rate flows where both the period and the packet size can be different for each flow. In this case, it is again possible to derive a close approximation for the delay distribution [13]. Empirical evidence supports the intuition that this queue leads to lower virtual waiting times than an M/G/1 queue where the service time distribution is the same as that of an arbitrary customer in the superposition. This is illustrated in Figure 4 for a particular traffic mix. As above we can further affirm that, in the strong variability ordering, each flow is better than a Poisson process of same size packets and that the superposition is therefore better than the M/G process.

A less precise bound is provided by assuming a Poisson arrival process of maximum (MTU) sized packets. The virtual waiting time distribution for the same load $M/D_{MTU}/1$ queue is also shown in Figure 4. The bound is clearly not as tight as that provided by M/G/1. However, it proves necessary to assume maximum sized packets to obtain a robust result with respect to the impact of jitter. Indeed, the NJ conjecture relies on the intuition that transforming a process by increasing the packet size (while maintaining other essential properties such as the mean intensity) worsens performance thus leading to upper bounds on performance.

Consider, therefore, the following transformation of a stationary point process of unit-sized packets denoted $N^{(1)}$: starting from an arbitrarily chosen packet, remove the next $m-1$ packets and replace the m^{th} by a "super-packet" of size m ; continue

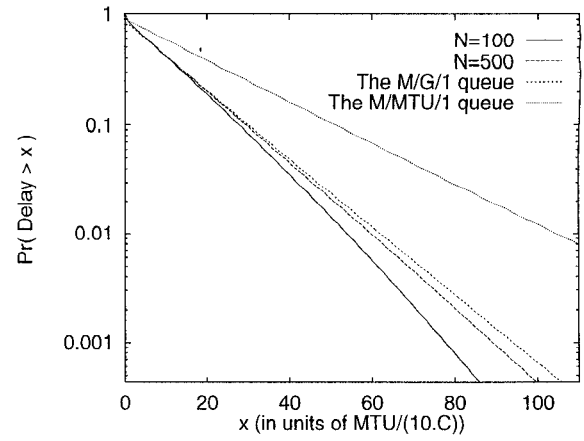


Fig. 4. Virtual waiting time distribution of the $\sum D_i^{X_i}/D/1$ queue compared with that of the corresponding M/G/1 and $M/D_{MTU}/1$ queues. Load $\rho = 0.8$. Superpositions of N periodic sources with packets of size $MTU/10$ and N sources with MTU-sized packets, all sources contributing the same load.

thus for all subsequent groups of m arrivals creating the process $N^{(m)}$. Intuitively, $N^{(1)}$ is better than $N^{(m)}$ in the sense that it produces smaller queues. It seems difficult, however, to prove that this is true in the ideal virtual waiting time ordering. It is proved in the weaker waiting time convex ordering in [1], Theorem 2.5.2, i.e., we have $N^{(1)} \leq_{wcx} N^{(m)}$. This allows us to state that the moments of the virtual waiting time process for $N^{(m)}$ are greater than those for $N^{(1)}$. In addition we have the following proposition:

Proposition 1: With the process $N^{(m)}$ constructed from a stationary process $N^{(1)}$ as described above, we have: $N^{(1)} \leq_{svar} N^{(m)}$.

Proof: Let $A_t^{(1)}$ and $A_t^{(m)}$ be the work generated during an arbitrary interval of length t by process $N^{(1)}$ and $N^{(m)}$, respectively. Let $\Pr[A_t^{(1)} = n] = \pi_n(t)$ and write $n = qm + r$ with $0 \leq r < m$. If one of the first r arrival epochs of $N^{(1)}$ in the interval is also an arrival epoch for $N^{(m)}$, we have $A_t^{(m)} = (q+1)m = n - r + m$. By stationarity, this occurs with probability r/m . Otherwise, $A_t^{(m)} = n - r$. Thus, for a convex function h ,

$$\begin{aligned} E[h(A_t^{(m)})] &= \sum_{n \geq 0} \left\{ \frac{r}{m} h(n - r + m) + \left(1 - \frac{r}{m}\right) h(n - r) \right\} \pi_n(t) \\ &\geq \sum_{n \geq 0} h(n) \pi_n(t) = E[h(A_t^{(1)})]. \end{aligned}$$

■

The strong variability ordering extends naturally to superpositions, as noted in Definition 3, i.e., we have $M + N^{(1)} \leq_{svar} M + N^{(m)}$ for any independent process M . The following proposition establishes that waiting time convex ordering is also preserved in a superposition.

Proposition 2: With process $N^{(m)}$ constructed from a stationary process $N^{(1)}$ as described above and M an arbitrary stationary marked point process, we have: $M + N^{(1)} \leq_{wcx} M + N^{(m)}$.

Proof: (outline) The proof is adapted from the proof that $N^{(1)} \leq_{wcx} N^{(m)}$ given in [1], page 221. For the process $M + N^{(m)}$ we retain all the arrival epochs of $M + N^{(1)}$ with

the mark ξ_i associated with arrival epoch i being either the same in both processes, if it corresponds to an arrival of M , or m , with probability $1/m$, and 0, with probability $(m-1)/m$, if it is an arrival of $N^{(m)}$. It is first necessary to show that we have convex ordering for the vector of marks (packet sizes) for n successive arrivals in the superposed process, i.e., $(\xi_1^{(m)}, \dots, \xi_n^{(m)}) \geq_{cx} (\xi_1^{(1)}, \dots, \xi_n^{(1)})$, for $n \geq 1$. This follows directly from the construction of process $N^{(m)}$ and, as in [1], allows the assertion that the workload at arrival epochs satisfies the increasing convex ordering. To extend this ordering to the stationary workload process, we can apply result (4.3.8) in [1] (page 238). ■

The above propositions allow us to affirm that, if m unit-sized packets constitute an MTU and if $N^{(m)}$ is better than Poisson/MTU, then $N^{(1)}$ is also better than Poisson/MTU in the waiting time convex and strong variability orderings.

V. ACCUMULATION OF JITTER

Constant rate streams lose their initial periodicity as they acquire jitter in multiplexer queues. It is important to be able to characterize this deformation and to evaluate its impact on queuing at nodes within the network. To illustrate the way jitter accumulates and to assess its impact we first restrict attention to the case of a simple tandem network with cross traffic. A test stream of constant sized packets goes through a series of multiplexing stages where it is in competition in FIFO queues with independent cross traffic (see Figure 5).

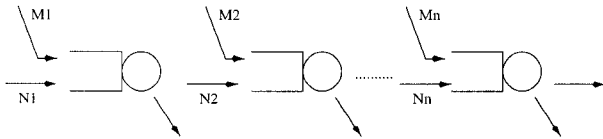


Fig. 5. Tandem queue with cross traffic

A. The saturated tandem

The tandem queue model can be analyzed exactly in the following case. Cross traffic consists of Poisson arrivals of packets of independent and identically distributed size and the intensity of the process at each queue is such that the server load is 1. This analysis generalizes the results derived for ATM traffic in [12] and [13]. Since the links are saturated, it is not necessary here to take explicit account of non-priority traffic.

Let the size of test flow packets be σ and the mean inter-arrival time be T . For simplicity, assume all queues in the tandem have the same unit service rate and the same cross traffic process. The latter is characterized by the arrival rate λ and the distribution of the packet size Σ . Since each queue is loaded to 1, we have: $\sigma/T + \lambda E[\Sigma] = 1$.

Let $(\tau_n^k)_{n \in \mathbb{Z}}$ denote the inter-arrival times of the test flow at the k^{th} queue. Every τ_n^k for fixed n is Markovian in k and evolves through the relation: $\tau_n^{k+1} = \sigma + A(\tau_n^k)$, where $A(t)$ denotes the cross traffic arriving in an interval of length t . This implies, as for ATM traffic [13], that if the test process is initially a renewal process it remains so at every stage and that a general test process converges to a renewal process as it progresses through more and more stages. It is possible to derive the moments of the inter-arrival distribution at each stage. In particular,

$E[\tau_n^k] = T$ for all k and $\text{Var}[\tau_n^{k+1}] = \lambda T E[\Sigma^2] + (1 - \sigma/T) \text{Var}[\tau_n^k]$. It follows from the latter relation that the inter-arrival time variance converges monotonically to $\text{Var}(\tau_n^\infty)$, where

$$\text{Var}(\tau_n^\infty) = \frac{(T - \sigma)}{1 - (1 - \sigma/T)^2} \frac{E[\Sigma^2]}{E[\Sigma]}.$$

If the test flow is initially periodic, its variance increases stage by stage towards this value. Given that $\Sigma \leq \text{MTU}$, the above expression is maximal when Σ is constant and equal to MTU. In other words, the test flow acquires maximal jitter (as measured by the inter-arrival time variance) when the cross traffic is Poisson/MTU. The squared coefficient of variation of the limit inter-arrival time distribution is then:

$$\text{CV}^2 = \frac{1 - \sigma/T}{1 - (1 - \sigma/T)^2} \frac{\text{MTU}}{T}.$$

We see that, for a constant test flow load σ/T , CV^2 increases as σ decreases. For small enough packets, the test flow is clearly not better than Poisson. However, assume $\sigma = \text{MTU}/m$ for m an integer and consider the squared coefficient of variation of the interval between the first and last of $m+1$ successive packets. This is independent of m and equal to $(1 - \sigma/T)/(2 - \sigma/T)$ which varies between 1/2 and 0 as the proportion of end-to-end traffic increases from 0 to 100%.

Now, by Propositions 1 and 2, this process is better, in the waiting time convex and strong variability orderings, than a process constructed by concentrating m successive packets on every m^{th} arrival epoch. The latter is a renewal process which has exactly the same properties as the test flow in the tandem with constant service time considered in [11]. We can thus conclude from Propositions 3 and 4 in that paper that, if this process is initially better than Poisson/MTU in the strong variability ordering, it remains so at each stage. By transitivity the initial flow of size σ packets is also better than Poisson/MTU in this ordering.

The same arguments can be applied when the servers in the tandem queue have different rates and when some cross flows remain in the tandem for more than one stage. The “better than Poisson/MTU” comparison still applies. It was shown above that cross flow packets smaller than MTU induce less jitter in the through flow. It also seems clear that a cross flow arrival process which is better than Poisson would also reduce jitter in the test flow.

B. Non-saturated networks

As for the case of ATM, to evaluate jitter accumulation in a non-saturated network it is necessary to turn to simulation. We have performed a number of experiments for the tandem network of Figure 5 with server load less than 1. None of the results obtained contradicts the NJ conjecture. Indeed, the evidence confirms our intuition that jitter in the test flow is reduced as load decreases.

Figures 6 and 7 relate to a tandem with Poisson/MTU cross traffic and through traffic consisting of an initially periodic flow of packets of size $\text{MTU}/10$ contributing a load of 0.03 at each stage. Note that in these simulations, when no EF packet is present, the link is immediately occupied for the transmission of non-priority packets of size MTU.

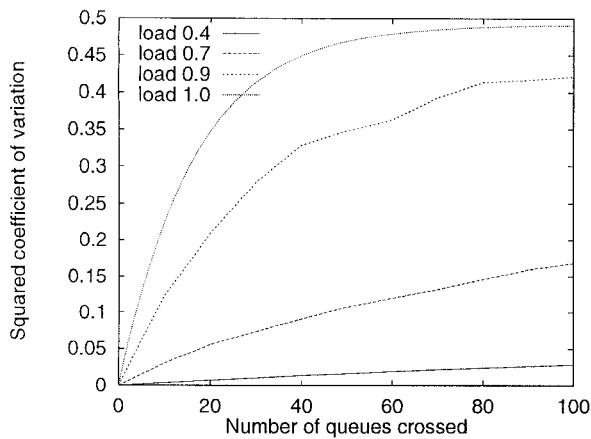


Fig. 6. Squared coefficient of variation of 10 successive intervals ($\sigma = \text{MTU}/10$).

Figure 6 shows the squared coefficient of variation of 10 successive inter-arrivals as a function of the number of stages crossed for server loads of 1, 0.9, 0.7 and 0.4. Recall from section IV-C that the original packet process is better than the process formed by concentrating arrivals on every 10th packet. The results in Figure 6 measure the deviation of the transformed through traffic from a Poisson/MTU stream for which the coefficient of variation is 1. As shown in the previous section, this coefficient increases from 0 to 0.5 with the number of links crossed when the load of every link is equal to 1. The simulation results confirm the intuition that variability is even less at lower loads. The variance of the 10-packet interarrival time is very small when load is less than 0.5.

The coefficient of variation is a satisfactory measure of burstiness for the saturated tandem where the through traffic stream is a renewal process. To further characterize this stream when link loads are less than 1, we consider the correlation between successive 10-packet intervals. Figure 7 shows the correlation coefficient for loads 1, 0.9, 0.7 and 0.4. This coefficient is zero for the renewal process pertaining at load 1. The figure clearly shows that correlation is negative and decreases with decreasing load. Negative correlation implies that the arrival processes are rather better than a renewal process with the same interarrival time distribution. This is because any short interval leading to a momentary queue is likely to be followed by a longer than average interval allowing this queue to be absorbed immediately.

Further simulations with cross traffic which is better than Poisson/MTU (smaller packets, smoother arrival process) confirm the intuition that this is a worse case for creating test flow jitter. In Figure 8 we show how the coefficient of variation of the 10-packet interval varies with the number of links for three different equal intensity cross traffics: Poisson/MTU, Poisson arrivals of packets whose size is uniformly distributed between $\text{MTU}/10$ and MTU , Erlang-2 arrivals (i.e., a renewal process with interpacket intervals equal to the sum of two exponential variables) of MTU sized packets. As postulated, the worst performance is indeed obtained for Poisson/MTU cross traffic.

The results presented in this section clearly do not prove the NJ conjecture. They do, however, constitute quite compelling evidence for its validity demonstrating that the jitter is reduced

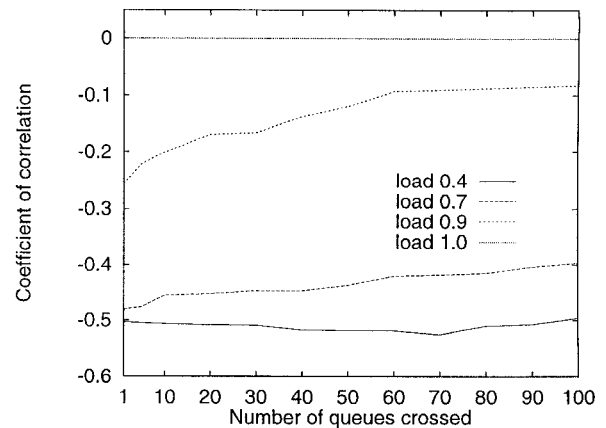


Fig. 7. Correlation coefficient for adjacent groups of 10 intervals ($\sigma = \text{MTU}/10$).

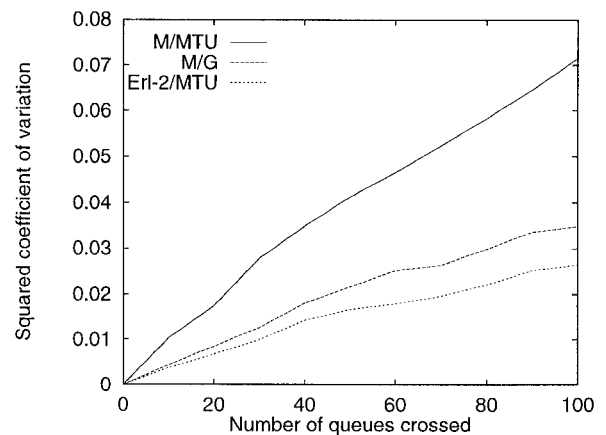


Fig. 8. Squared coefficient of variation of 10 successive intervals with different cross traffics: Poisson and Erlang-2 arrivals of MTU -sized packets, Poisson arrivals of packets of size uniformly distributed between $\text{MTU}/10$ and MTU . Load $\rho = 0.5$.

and performance enhanced as we deviate from the previously analysed worst case saturated tandem.

C. The large deviations ordering

The sample path large deviations ordering introduced in [11] for constant packets can also be applied in the present case to derive “better than” comparisons in a more general network setting than that of the tandem queue. Instead of point processes, we must focus on the amount of work generated by individual flows. Let $A(0, t]$ denote the amount of work generated in an interval of length t , and define:

$$A^n(t) = \begin{cases} \frac{1}{n} A(0, nt] & \text{if } t \geq 0, \\ -\frac{1}{n} A(nt, 0] & \text{if } t < 0. \end{cases}$$

The function $t \mapsto A^n(t)$ belongs to the space $D(\mathbb{R}, \mathbb{R})$ of right continuous functions with left-hand limits. We say that the sequence A^n satisfies a sample path Large Deviations Principle (sp-LDP) if it satisfies a Large Deviations Principle in the space $D(\mathbb{R}, \mathbb{R})$ with a rate function $I_A : D(\mathbb{R}, \mathbb{R}) \mapsto \mathbb{R}$. For instance, for a Poisson process M of arrivals of unit-sized packets, the sequence M^n satisfies an sp-LDP and the corresponding rate

function is given by:

$$I_M(\gamma) = \begin{cases} \int_{-\infty}^{+\infty} H_M(\gamma'(t)) dt, & \text{if } \gamma \text{ is absolutely} \\ & \text{continuous and } \gamma(0) = 0, \\ +\infty, & \text{otherwise} \end{cases}$$

where

$$H_M(x) = \begin{cases} x \ln \frac{x}{\lambda} - x + \lambda & \text{if } x \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

For details, see [11]. We define the LD ordering as follows.

Definition 5: Large deviations ordering. $A \leq_{LD} A'$ if and only if, $I_A(\gamma) \geq I_{A'}(\gamma)$, for $\gamma \in D(\mathbb{R}, \mathbb{R})$.

Theorem 1: Consider a feed-forward network of FIFO queues as defined in [11]. Assume stability, i.e., at each queue the load is strictly less than 1. If flows are better than Poisson/MTU for the LD ordering at their origin, they retain this property throughout the network. As a consequence, if $V(\{Poisson, MTU\}, c)$ denotes the stationary workload of an arbitrary server when the real traffic A is replaced by a Poisson process of arrivals of MTU-sized packets, we have:

$$V(A, c) \leq_{EB} V(\{Poisson, MTU\}, c)$$

Proof: The proof can be derived directly from that in [11] on replacing the arrival point process $N(0, t]$ by work process $A(0, t]$. ■

This theorem constitutes a demonstration that the NJ conjecture is true in an asymptotic regime, at least for the class of feed-forward networks.

VI. ACCOUNTING FOR VARIABLE RATE FLOWS

Audio and video flows requiring the low transfer delays provided by EF are generally not CBR. Variable rate video coding in particular produces extreme rate variations at multiple time scales. Multiplexing such flows with statistical performance guarantees can be performed with EF using so-called “rate envelope multiplexing”.

A. Rate envelope multiplexing

Assume for the moment that the rate of a flow i is a well defined quantity $\lambda_i(t)$ which varies slowly except at the epochs of a low intensity point process representing discontinuous rate changes (e.g., a flow changing state from “on” to “off”). Rate envelope multiplexing consists in ensuring that, for the set $\mathcal{F}$ of flows sharing a link of rate C , we have $Pr[\sum_{i \in \mathcal{F}} \lambda_i(t) > C] < \epsilon$, for some suitably small number ϵ (see Fig. 9).

Rate envelope multiplexing is “bufferless” in a hypothetical fluid system. Delay is then null and loss can be limited by applying admission control to guarantee a sufficiently small rate overload probability ϵ . Bufferless multiplexing is efficient when the rate of any flow is never more than a small fraction of the service rate C . The performance of bufferless multiplexing is discussed in [13], Chapter 16, for instance. Performance of the actual non-fluid system depends additionally on the dynamics of the queue necessary to take account of coincident packet arrivals from distinct flows.

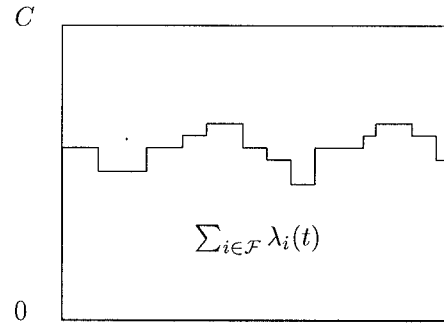


Fig. 9. Rate envelope multiplexing: the combined instantaneous rate of multiplexed flows only exceeds the service rate C with small probability ϵ

B. Negligible jitter

Assume flows are shaped at the network ingress to their nominal peak rate p_i , for $i \in \mathcal{F}$. They may thus be considered to have an on/off traffic pattern, an off-period beginning whenever the inter-packet interval of the shaped stream exceeds σ_i/p_i , where σ_i is the constant size of flow i packets. Figure 10 depicts the imagined composition of on- and off-periods for three flows. Note that an isolated packet of size s (occurring when the actual flow rate is less than the nominal rate p , for instance) gives rise to an on-period of duration s/p . We consider the packets to remain in the same on-period throughout the network even though the inter-packet spacing changes at each queuing stage.

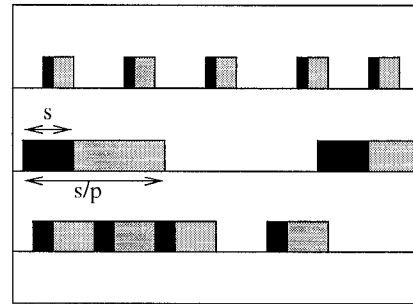


Fig. 10. EF flows shaped to a common rate p . The figure shows three flows with packets of different sizes represented by the black rectangles. The on-periods are represented as a contiguous sequence of black and gray rectangles.

If the sum of the peak rates of active flows were to be always less than the service rate, the NJ conjecture suggests that jitter in the initially shaped flows remains negligible. In other words, queuing delays can be bounded as in Section III-B.2 by assuming flows correspond to a Poisson/MTU process. In practice, it must be ensured that the probability of exceeding the tolerable utilization level derived from the M/D/1 model is small enough. It is additionally necessary to avoid the accumulation of delays whenever rate overload does occur. A possible rule of thumb would be to limit the queue length to the time to send an MTU worth of data at the rate of the highest rate flow multiplied by the service rate C .

C. Admission control

An extensive literature exists on the subject of admission control for rate envelope multiplexing. The most satisfactory ap-

proach is to rely on a minimal description of each flow coupled with a measurement-based estimation of current traffic, as in [5], for example. However, even if the only traffic descriptor is the flow peak rate p_i , the requirement to signal this rate and to maintain per flow state for a potentially large number of streaming flows may lead to scalability problems.

One possible approach we are currently exploring is to attribute the same peak rate p^* to all streaming flows. This rate defines a maximum peak rate for any flow, i.e., $p^* \geq p_i$ for $i \in \mathcal{F}$. In the interest of multiplexing efficiency, p^* must be a small fraction of the link rate (1%, say). Flows shaped to this common rate would appear to the network as on/off flows and, provided an adequate measurement-based admission control were implemented, would conserve negligible jitter throughout the network. We envisage that new flows would be identified on the fly without signaling and rejected, if necessary, simply by discarding their packets. The choice of a common rate avoids the need for signaling and reduces required state to a list of flow identifiers. This list is necessary to be able to recognize the appearance of packets from a new flow.

VII. CONCLUSION

Reliance on deterministic performance bounds for packet delay within EF aggregates can be extremely pessimistic. It also requires complex traffic management rules to ensure the set of established flows within a network is admissible. On the other hand, statistical bounds are clearly sufficient for streaming applications if it can be guaranteed that packet delay and loss are small enough. It is our contention that such guarantees are feasible when EF is realized using priority queuing, flows are independent and their packets are spaced to a nominal peak rate at the network ingress.

In this case, it is sufficient to estimate delays and buffer saturation probabilities assuming EF traffic at any queue in the network can be assimilated to a Poisson stream of MTU-sized packets. This is a necessary assumption if there is no limit on the number of (low rate) flows admitted and no control on packet size. Sufficiency relies on a conjecture that variable packet delays accumulating in successive queuing stages do not cause individual flows to become more bursty than a Poisson process of MTU packets. This is the negligible jitter (NJ) conjecture.

We have shown that, if the NJ conjecture is true, then bounds on end-to-end delay can be evaluated simply as the sum of waiting times in independent M/D/1 queues. Under realistic network assumptions, this bound is sufficiently small for real time interactive streaming applications. The M/D/1 model can also be used for buffer sizing. If the NJ conjecture is true, the statistical bounds are rigorous and not an approximation. The major contribution of the paper has been to investigate its validity by analysis and by simulation.

Analytical results on the comparison of stochastic processes show that the conjecture is true for the so-called strong variability ordering in the particular case of a quasi-saturated tandem network where a test flow is subjected to interference from Poisson cross traffic in successive stages. Further analytical comparisons and simulation results confirm that the characteristics of the test flow improve as the cross traffic assumptions are relaxed: lower load, arrival process smoother than Poisson,

smaller packets, flows interfering for more than one stage. The conjecture is also proved to be valid for "better than" comparisons in the asymptotic effective bandwidth ordering in the case of more general feed forward networks. These results generalize the negligible CDV conjecture which has already been formulated for ATM and used as the basis for B-ISDN traffic engineering standards.

Statistical guarantees based on the NJ conjecture can be applied to both constant and variable rate streaming traffic. We have outlined a possible approach to traffic management for variable rate EF traffic using measurement-based admission control which avoids the need to know individual flow traffic parameters at routers beyond the network edge.

The NJ conjecture provides a framework for understanding the results of previous simulation studies which demonstrated the limited impact of jitter accumulation on packet delay. While a complete formal proof of the conjecture appears beyond reach, the evidence presented here strongly supports the intuition that it is true. In the absence of any obvious counter examples, we believe that the NJ conjecture and the engineering rules which are derived from it can be used to provide practically useful statistical guarantees for streaming applications in the Internet.

POST SCRIPTUM It has recently been pointed out at the IETF that the definition of EF behavior in [10] is flawed. This does not invalidate the present analysis, however, which is based on the common understanding that, whatever its eventual formal definition, EF can be realized using priority queuing.

REFERENCES

- [1] F. Baccelli, P. Brémaud. Elements of Queuing Theory. Springer Verlag, 1994.
- [2] F. Brichet, L. Massoulié and J. Roberts. Stochastic ordering and the notion of negligible CDV. *Proceedings of ITC 15*, Elsevier, 1997.
- [3] A. Charny, J.-Y. Le Boudec. Delay bounds in a network with aggregate scheduling. *QoSIS Workshop, Berlin*, 2000.
- [4] N.G. Duffield and N. O'Connell. Large Deviations and overflow probabilities for the general single server queue with applications. *Proc.Camb. Phil. Soc.*, 188, 1995.
- [5] R. Gibbens, F. Kelly, P. Key. A decision theoretic approach to call admission control in ATM networks. *IEEE JSAC*, Vol 13, N°6, pp 1101-1114, August 1995.
- [6] P. Goyal, A. Greenberg, C. Kalmanek, W. Marshall, P. Mishra, D. Nortz and K. Ramakrishnan. Integration of call signaling and resource management for IP telephony. *IEEE Network*, pp 24-32, May/June 1999.
- [7] M. Grossglauser and S. Keshav. On CBR service. In *Proceedings of INFOCOM96*, pages 129-137, San Francisco, CA, March 1996.
- [8] R. Guérin, V. Pla. Aggregation and Conformance in Differentiated Service Networks: A Case Study. *ITC Seminar on IP Traffic*, Monterey, September 2000.
- [9] ITU Recommendation E.736. Methods for cell level traffic control in B-ISDN. ITU, Geneva, 2000.
- [10] V. Jacobson, K. Nichols, K. Poduri. An Expedited Forwarding PHB. *RFC 2598*, June 1999.
- [11] L. Massoulié. Large Deviations Ordering of Point Processes in some Queuing Networks. *Queuing Systems*, 1999.
- [12] W. Matragi, K. Sohraby and C. Bidiskian. Jitter calculus in ATM networks, Multiple nodes. *IEEE/ACM Trans on Networking*, Vol 5, No 1, pp122-133, Feb. 1997.
- [13] J. Roberts, U. Mocci, J. Virtamo (Eds). Broadband Network Teletraffic (Final Report of COST 242). *LNCS 1155 Springer Verlag*, 1996.
- [14] J. Sahni, P. Goyal, and H. M. Vin. Scheduling CBR flows: FIFO or per-flow queuing? In *Proceedings of NOSS-DAV99*, AT&T Learning Center, Basking Ridge, NJ, June 1999.
- [15] R. Szekli. Stochastic Ordering and Dependence in Applied Probability. Springer Verlag, 1995.