



# Finite buffer queue with generalized processor sharing and heavy-tailed input processes <sup>☆</sup>

Predrag Jelenković <sup>\*</sup>, Petar Momčilović

*Department of Electrical Engineering, Columbia University, 500 West 120th Street, New York, NY 10027, USA*

---

## Abstract

We consider a set of  $N$  fluid On–Off flows that share a common server of capacity  $c$  and a finite buffer  $B$ . The server capacity is allocated using the generalized processor sharing scheduling discipline. Each flow has a minimum service rate guarantee that exceeds its long-term average demand  $\rho_i$ . The buffer sharing is unrestricted as long as there is available space. If the buffer is full, the necessary amount of fluid from the most demanding flows is discarded. When On periods are heavy-tailed, we show that the loss rate of a particular flow  $i$  is asymptotically equal to the loss rate in a reduced system with capacity  $c - \sum_{j \neq i} \rho_j$  and buffer  $B$ , where this flow is served in isolation. In particular, the system behaves as if it had  $N$  times bigger buffer. This new insight on buffer multiplexing gain offers an additional tradeoff in distributing buffers between core and edge switching elements.

© 2002 Elsevier Science B.V. All rights reserved.

*Keywords:* Network multiplexer; Long-range dependence; Finite buffer fluid flow

---

## 1. Introduction

Modern communication networks are engineered to carry a diverse spectrum of multimedia services, ranging from real-time traffic, such as voice and video, to various data- and Web-related applications. These services have different quality of service (QoS) requirements, e.g., real-time services have stringent delay requirements, but can tolerate relatively high losses. On the other hand, data-related services typically could tolerate larger

delays, but need minimal or no losses. Providing the QoS differentiation in integrated multimedia networks is usually achieved through priority scheduling mechanisms. The most popular scheduling schemes, e.g., weighted fair queueing, are based on the generalized processor sharing (GPS) algorithm. These algorithms offer the flexibility for providing a high degree of service differentiation, extracting statistical multiplexing gains as well as protecting individual flows from the ones with high service demands.

Rigorous investigation of stochastic systems with GPS dates back to [9]; see also [9] for some earlier references. This work was centered around the problem of time-shared computer systems. Recent results for traffic models with exponential characteristics can be found in [1,18,20,21]. However, comprehensive statistical measurements in

---

<sup>☆</sup> This work is supported by the National Science Foundation Presidential Early Career Accomplishment Award No. 9875156.

<sup>\*</sup> Corresponding author.

*E-mail addresses:* [predrag@ee.columbia.edu](mailto:predrag@ee.columbia.edu) (P. Jelenković), [petar@ee.columbia.edu](mailto:petar@ee.columbia.edu) (P. Momčilović).

currently deployed networks have found repeated evidence of the presence of a high degree of statistical dependency and variability, often referred to as self-similarity, in network traffic streams, e.g., see [7,10,15,16,19]. The most common approaches to modelling these intricate statistical phenomena are through long-range dependent Gaussian processes, such as fractional Brownian motion, and heavy-tailed semi-Markov type processes, e.g., On–Off processes. In this paper we focus on the latter.

Motivated by these empirical findings, a series of research studies that develop new analytical techniques for evaluating stochastic heavy-tailed queueing systems have appeared. In particular, very recent investigations of the behavior of GPS in the presence of heavy-tailed arrival streams can be found in [2–6]. The reader may consult the same papers for additional references on GPS and heavy tails. These papers consider a system with a finite number of heavy-tailed sessions each of which is queued in an infinite buffer queue; the content of the queues is served by a single server that is scheduled using GPS. When the GPS weights are appropriately chosen, these studies show that each session experiences the same queueing behavior as if it were served in isolation with an appropriate constant capacity. However, if the weights are not properly engineered, the flows may experience induced burstiness [3,6]; these results stress the importance of properly selecting the GPS weights.

In this paper we investigate a queueing system in which the heavy-tailed On–Off sessions share a single finite buffer  $B$ . The buffer sharing is unrestricted as long as there is available space; if the buffer is full, the necessary amount of fluid from the most demanding flows is discarded. The flows share a common server according to the GPS scheduling discipline. A formal description of our model and some preliminary results are presented in Sections 2 and 3, respectively. When each flow receives the minimum service rate guarantee that exceeds its long-term average demand  $\rho_i$ , our main result, stated in Theorem 4.1 of Section 4, shows that the loss rate of a particular flow  $i$  is asymptotically equal to the loss rate in a reduced system with capacity  $c - \sum_{j \neq i} \rho_j$  and buffer  $B$ , where this flow is served in isolation. In other words, each

flow perceives having the entire buffer to itself. This finding complements the result derived in [2] by quantifying the benefits of having a shared buffer. The new qualitative insight may prove useful in deciding on whether to engineer buffers at the periphery or in the core of the network. Section 5 contains numerical examples that illustrate an excellent agreement between the analytical and simulation results. It is worth observing that the derived asymptotic approximation yields accurate results even for relatively large probabilities in the range of  $10^{-2}$ – $10^{-3}$ . Concluding remarks and a brief discussion of engineering implications are presented in Section 6.

## 2. Model description

Consider  $N$  independent On–Off flows  $A_i(t)$ ,  $1 \leq i \leq N$ . An On–Off flow consists of an alternating sequence of independent activity and silence periods. Activity and silence periods of the  $i$ th flow are equal in distribution to  $\tau_i$  and  $\eta_i$ , respectively. Assume that  $\tau_i$  and  $\eta_i$  are almost surely (a.s.) positive and that their expectations  $\mathbb{E}\tau_i$  and  $\mathbb{E}\eta_i$  are finite. During an activity period flow  $i$  generates fluid at constant rate  $A_i(t) = r_i$  and for a period of silence  $A_i(t) = 0$ . Suppose that  $A_i(t)$  is right continuous and stationary; for a precise construction see [11]. Then, the stationary probability that flow  $i$  is active and its long-term average rate are equal to  $p_i = \mathbb{P}[A_i(t) = r_i] = \mathbb{E}\tau_i / (\mathbb{E}\tau_i + \mathbb{E}\eta_i)$  and  $\rho_i = \mathbb{E}A_i(t) = p_i r_i$ , respectively.

These On–Off flows share a common server of capacity  $c$  and a buffer space of size  $B$ . Let  $W_i(t)$  be the unfinished work of flow  $i$  at time  $t$ ,  $\sum W_i(t) \leq B$ . The server capacity is distributed among flows according to the GPS scheme. Each flow  $i$  is assigned a weight  $\phi_i > 0$  such that  $\sum_{i=1}^N \phi_i = 1$ . Weight  $\phi_i$  represents the guaranteed share of the server capacity for flow  $i$ . Available excess capacity is redistributed among flows according to the GPS weights  $\phi_i$ . Service rates  $c_i(t)$  for each flow  $i$  at time  $t$  can be computed by using a recursive algorithm described in [8, pp. 4–5]. Let  $E(t)$  be the set of flows with  $W_i(t) = 0$ , that are receiving service at rate  $c_i(t) = A_i(t)$ , i.e.,  $E(t) \triangleq \{i: W_i(t) = 0, c_i(t) = A_i(t)\}$ . Then, it is not difficult

to see that  $E(t) = \{i: W_i(t) = 0\}$  almost everywhere (a.e.) Lebesgue. Therefore, using this property and the characteristics of GPS for any flow  $i \notin E(t)$ , rate  $c_i(t)$  satisfies

$$c_i(t) = \frac{\phi_i \left( c - \sum_{j \in E(t)} A_j(t) \right)}{\sum_{j \notin E(t)} \phi_j} \geq c \phi_i \text{ a.e.} \quad (2.1)$$

Buffer sharing is unrestricted as long as there is available space, i.e., the workloads evolve as if they were in the infinite buffer system. When the buffer fills up, the flows with the maximum amount of fluid  $W_i(t)$  in the buffer are subject to penalty. They will experience the minimum necessary loss of fluid such that the flows with smaller workloads can be accommodated. We exemplify this policy for two On–Off flows with equal GPS weights in Fig. 1. Possible extensions to more general buffer sharing policies will be briefly discussed at the end of Section 4.

More formally, following the approach from [8], the evolution of  $W_i$ -s can be described with a set of differential equations. In order to account for the finiteness of the buffer space we define

$$D(t) = \left\{ i: W_i(t) = \max_{1 \leq j \leq N} W_j(t), \sum_{j=1}^N W_j(t) = B \right\}$$

with  $|D(t)|$  denoting the cardinality of  $D(t)$ ; note that  $D(t)$  is nonempty only if the buffer is full. The elements of  $D(t)$  are flows that could potentially experience losses at time  $t$ . Let  $M(t)$  be the largest subset of  $D(t)$  such that for all  $i \in M(t)$  the following inequality holds:

$$A_i(t) - c_i(t) > -|M(t)|^{-1} \sum_{j \notin M(t)} (A_j(t) - c_j(t)).$$

The workloads of the flows in  $M(t)$  are regulated according to

$$\dot{W}_i(t) = -|M(t)|^{-1} \sum_{j \notin M(t)} (A_j(t) - c_j(t));$$

this adjustment is necessary to ensure  $\sum W_i = B$ . Note that  $\dot{W}_i(t)$  can be both positive and negative. For example, consider a system of three processes with equal GPS weights. Let  $W_1(0) = W_2(0) = 2B/5$ ,  $W_3(0) = B/5$ ,  $A_1(0) = A_2(0) = c$  and  $A_3(0) = 0$ . Then,  $M = \{1, 2\}$  and, hence,  $\dot{W}_1(0) = \dot{W}_2(0) = c/6$ , i.e., the workloads of the first two processes

are increasing at rate  $c/6$ . However, without any adjustment the workloads would increase at rate  $2c/3$ .

Next, similarly as for the set  $E(t)$ , we observe that  $M(t) = D(t)$  a.e. (Lebesgue). Thus, the evolution of the individual workloads is a.e. described by the following set of differential equations:

$$\dot{W}_i(t) = \begin{cases} 0 & \text{if } i \in E(t), \\ -|D(t)|^{-1} \sum_{j \notin D(t)} (A_j(t) - c_j(t)) & \text{if } i \in D(t), \\ A_i(t) - c_i(t) & \text{otherwise.} \end{cases} \quad (2.2)$$

The functions  $W_i(t)$  are absolutely continuous and therefore sets of Lebesgue measure zero can be

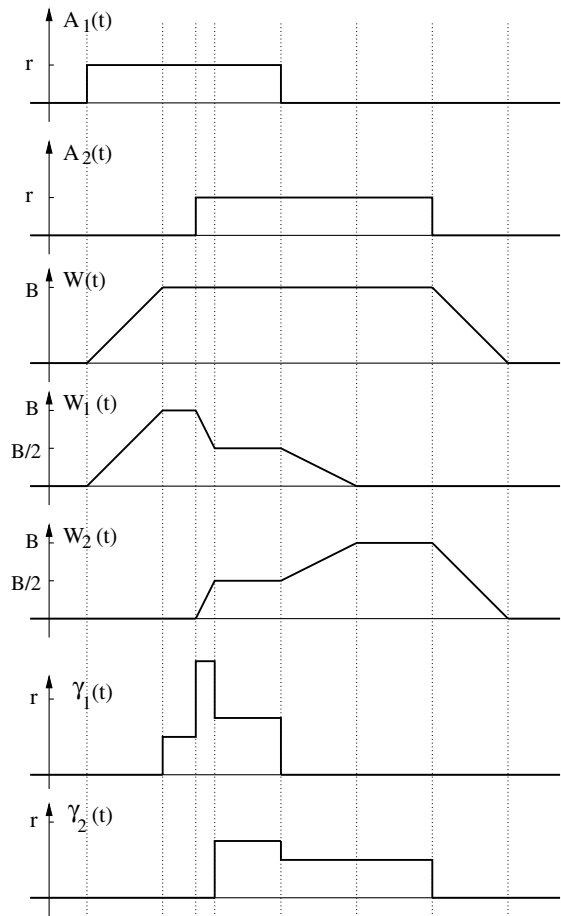


Fig. 1. An illustration of the buffer management policy. Two On–Off flows with peak rate  $r$  share a server of capacity  $c = r/2$  and buffer of size  $B$ . The GPS weights of the two flows are equal,  $\phi_1 = \phi_2 = 0.5$ .

ignored for purpose of their characterization [8]. Hence, Eq. (2.2) completely define the behavior of the system. We assume that this system of equations has a unique stationary solution and, unless otherwise specified,  $W_i(t)$  will be used to denote this solution. Let  $A_i$ ,  $c_i$ ,  $\bar{W}_i$  be random variables equal in distribution to  $A_i(t)$ ,  $c_i(t)$ ,  $\bar{W}_i(t)$  in stationarity, respectively.

The main objective of this paper is the asymptotic computation of the long-term average loss rate for a given flow. We say that flow  $i$  is overflowing at time  $t$  if  $A_i(t) - c_i(t) > \bar{W}_i(t)$ . The instantaneous loss rate process for flow  $i$  is defined as

$$\gamma_i^B(t) \triangleq A_i(t) - c_i(t) - \bar{W}_i(t), \quad (2.3)$$

and its expected loss rate is equal to  $\Gamma_i^B = \mathbb{E}\gamma_i^B(t)$ . Note that the buffer management policy implies

$$\begin{aligned} & \{A_i(t) - c_i(t) > \bar{W}_i(t)\} \\ & \subseteq \left\{ W_i(t) \geq B/N, \sum_{i=1}^N W_i(t) = B \right\}. \end{aligned} \quad (2.4)$$

This inclusion is a simple consequence of the fact that if the buffer is full and  $W_i(t) < B/N$ , then there exists a flow  $j$  with higher workload than that of flow  $i$ , and, therefore, flow  $i$  cannot experience losses. It is worth mentioning that the described queueing system is work- and buffer-conserving. Work-conservation follows from the properties of GPS; buffer-conservation means that no flow will experience loss of fluid unless the buffer is full.

### 3. Preliminary results

This section contains preliminary sample path bounds and some of the existing results from the literature on fluid queues with heavy-tailed On–Off flows that will be used in deriving our main results.

#### 3.1. Sample path bounds

We start with an introduction of a finite buffer queueing process that will be used to bound the workload processes in the GPS system with buffer sharing that we have described in Section 2. Consider a fluid queue with constant capacity  $c$ , finite buffer  $B$  and arrival process  $A(t)$ . Informally,

at time  $t$ , fluid arrives to this system at rate  $A(t)$  and is leaving the system at rate  $c$ . When the queue level reaches the buffer limit  $B$  fluid arriving in excess of the draining rate  $c$  is lost. We use  $Q_A^{B,c}(t) \in [0, B]$  to denote the queue content at time  $t$ .

In this paper we only consider arrival processes  $A(t)$  that are right continuous piece-wise with a.s. increasing jump times  $\{T_0 = 0 < T_1 < T_2 < \dots\}$ . In this case, for any initial value  $Q_A^{B,c}(0)$  and  $t \in (T_n, T_{n+1}]$ ,  $n \geq 0$ , the evolution of  $Q_A^{B,c}(t)$  is given by

$$Q_A^{B,c}(t) = \min \left( (Q_A^{B,c}(T_n) + (t - T_n)(A(T_n) - c))^+, B \right), \quad (3.1)$$

where  $(x)^+ = \max(0, x)$ . When  $A(t)$ , i.e.,  $\{(T_{n+1} - T_n), A(T_n)\}$ , is stationary and ergodic, and  $\mathbb{E}A(t) < c$ , recursion (3.1) has a unique stationary and ergodic solution, see [17]. Let  $Q_A^{B,c}$  and  $A$  be random variables that are equal in distribution to  $Q_A^{B,c}(t)$  and  $A(t)$  in stationarity, respectively. The loss rate for the described system,  $A_A^{B,c}$  is defined as

$$A_A^{B,c} \triangleq \mathbb{E}[(A - c)\mathbf{1}\{Q_A^{B,c} = B\}].$$

The first sample path bound relates the amount of unfinished work of flow  $i$  in the GPS system and the workload of a queue in which flow  $i$  is served in isolation.

**Proposition 3.1.** *If  $W_i(t) \leq Q_{A_i}^{B, \phi_i c}(t)$  for  $t = 0$ , then the inequality holds for all  $t \geq 0$ .*

**Proof.** Follows from inequality (2.1) and the fact that the portion of the buffer available to  $W_i(t)$  is not greater than  $B$ .  $\square$

In order to simplify the notation we set  $A(t) = \sum_{i=1}^N A_i(t)$ ,  $r = \sum_{i=1}^N r_i$ ,  $A_{-i}(t) = \sum_{j \neq i} A_j(t)$  and  $\rho_{-i} = \sum_{j \neq i} \rho_j$ . The second sample path bound formalizes an intuitively expected notion that multiplexing reduces the aggregate queueing workload. The proof of the following proposition can be found in [11].

**Proposition 3.2.** *If  $Q_A^{B, \sum_{n=1}^N c_n}(t) \leq \sum_{n=1}^N Q_{A_n}^{B, c_n}(t)$  for  $t = 0$ ,  $c_n > 0$ , then the inequality holds for all  $t \geq 0$ .*

Here, we state another sample path bound from [11]. This bound will be used to limit the amount of free buffer space in a finite buffer queue. Let quantity  $Q_c^{\infty,A}(t)$  represent the workload in an infinite buffer fluid queue with constant arrival rate  $c$  and variable service rate  $A(t)$ .

**Proposition 3.3.** *If  $B - Q_A^{B,c}(t) \leq Q_c^{\infty,A}(t)$  for  $t = 0$ , then the inequality holds for all  $t \geq 0$ .*

### 3.2. Queueing results

This subsection consists of the known results on fluid queues with heavy-tailed On–Off flows that will be used in proving our main result. In the analysis of renewal processes excess (or residual) random variables and distribution functions play an important role. For a non-negative random variable  $X$  with distribution  $F$  and finite mean  $\mathbb{E}X$ , the excess distribution  $F^e$  is defined by  $F^e(x) = (\mathbb{E}X)^{-1} \int_0^x (1 - F(u)) du$ ,  $x \geq 0$ . A random variable  $X^e$  with d.f.  $F^e$  is called the excess variable of  $X$ . Throughout the paper, for any two real functions  $f(x)$  and  $g(x)$ , we use the frequently used notation  $f(x) \sim g(x)$  as  $x \rightarrow \infty$  to denote  $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$ .

The theorem below summarizes the results from [13,14] in the case when the arrival process is a single On–Off flow. The common heuristic explanation of this theorem comes from identifying the most likely buffer overflow scenario, which, in this system, is due to an isolated very long On period, see Fig. 2. Hence, in stationarity, the overflow of a large buffer occurs when the source is in its active state,  $\{A_i = r_i\}$ , and the current On period has lasted long enough to fill up the buffer, i.e.,  $\{\tau_i^e > B/(r_i - c)\}$ .

Classes of heavy-tailed distributions  $\mathcal{IR}$ ,  $\mathcal{I}$  ( $\mathcal{IR} \subset \mathcal{I}$ ) are defined in Appendix. Class  $\mathcal{IR}$

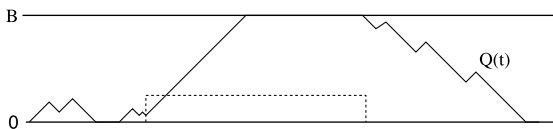


Fig. 2. The most likely overflow scenario in a finite buffer fluid queue with a single On–Off process. The long On period causing the overflow is shown with the dash line.

includes the Pareto family,  $\mathbb{P}[X > x] = (x - \beta + 1)^{-\alpha}$ ,  $x \geq \beta \geq 0$ ,  $\alpha > 0$ . In addition class  $\mathcal{I}$  includes the Weibull distribution,  $\mathbb{P}[X > x] = \exp(-x^\beta)$  for  $0 < \beta < 1$ , and the lognormal distribution  $\mathbb{P}[X \leq x] = \Phi((\log x - \mu)/\sigma)$ , where  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ ,  $\Phi$  is the standard normal distribution.

**Theorem 3.1.** *Let  $r_i > c > \rho_i$  and  $\tau_i^e \in \mathcal{I}$ . Then,  $A_{A_i}^{B,c} = (r_i - c)\mathbb{P}[Q_{A_i}^{B,c} = B]$  and as  $B \rightarrow \infty$*

$$\begin{aligned} \mathbb{P}[Q_{A_i}^{B,c} = B] &\sim \frac{c - \rho_i}{\rho_i} \frac{p_i}{1 - p_i} \mathbb{P}[Q_{A_i}^{\infty,c} > B] \\ &\sim p_i \mathbb{P}\left[\tau_i^e > \frac{B}{r_i - c}\right]. \end{aligned}$$

Furthermore, if  $\tau_i^e \in \mathcal{IR}$  then

$$\lim_{\epsilon \downarrow 1} \overline{\lim}_{B \rightarrow \infty} \frac{\mathbb{P}[Q_{A_i}^{B,c} \geq \epsilon B]}{\mathbb{P}[Q_{A_i}^{B,c} = B]} = 1.$$

The next theorem provides the asymptotic characterization of the aggregate loss rate in a finite buffer queue fed by  $N$  independent On–Off flows. Due to the work-conservation the result applies to the total workload  $\sum W_i$  in our GPS model. Here, the most likely buffer overflow arises when several very long On periods overlap. Typically, there is more than one way to overflow, which is accounted for with a summation in the theorem. For a detailed proof and experimental investigation of its accuracy see [11] and [12], respectively.

**Theorem 3.2.** *Let  $r_\mathcal{Q} \triangleq \sum_{j \in \mathcal{Q}} r_j + \sum_{j \notin \mathcal{Q}} \rho_j$  for all  $\mathcal{Q} \subseteq \{1, \dots, N\}$  and*

$$\mathcal{O} \triangleq \{\mathcal{Q} : r_\mathcal{Q} - (r_j - \rho_j) < c \leq r_\mathcal{Q}, \forall j \in \mathcal{Q}\}.$$

If  $\tau_i^e \in \mathcal{IR}$  for  $1 \leq i \leq N$  and  $r_\mathcal{Q} \neq c$  for all  $\mathcal{Q} \in \mathcal{O}$ , then as  $B \rightarrow \infty$

$$A_A^{B,c} \sim \sum_{\mathcal{Q} \in \mathcal{O}} (r_\mathcal{Q} - c) \prod_{j \in \mathcal{Q}} p_j \mathbb{P}\left[\tau_j^e > \frac{B}{r_\mathcal{Q} - c}\right].$$

In order to illustrate the preceding theorem we consider the following example. Let a server of capacity  $c = 7$  be shared by three On–Off flows

with peak rate  $r = 3$  and average rate  $\rho = 0.3$ . Then set  $\mathcal{O}$  consists of a single element,  $\mathcal{O} = \{1, 2, 3\}$ . If for all three flows the excess On periods are equal in distribution to  $\tau^c \in \mathcal{I}\mathcal{R}$ , then the loss rate is asymptotically equal to  $0.002 \mathbb{P}^3[\tau^c > 0.5B]$ .

#### 4. Main result

Finally, we are ready to state our main result. It establishes the buffer equivalence between the GPS system and a finite buffer queue in which a flow is served in isolation.

**Theorem 4.1.** *Let  $\phi_j c > \rho_j$  for all flows. If  $\tau_i^c \in \mathcal{I}\mathcal{R}$  and  $r_i > c - \rho_{-i}$  for some flow  $i$ , then as  $B \rightarrow \infty$*

$$\Gamma_i^B \sim A_{A_i}^{B,c-\rho_{-i}}.$$

**Heuristic explanation.** Since there is a strict stability  $\phi_j c > \rho_j$  for all flows, this ensures that the workload build-up of process  $i$  is unlikely to be caused by other flows. This fact and the heavy tailed nature of On periods of process  $i$  result in the most likely overflow scenario being due to a single long On period of process  $i$ . Therefore, during an overflow, with very high probability all other flows exhibit average behavior, while flow  $i$  transmits at its peak rate  $r_i$ . This implies that the buffer fills up at rate  $r_i + \rho_{-i} - c$ . In addition, the average behavior of processes  $j \neq i$  yields  $W_j = O(1)$ , and, thus, process  $i$  can potentially occupy up to  $B - O(1)$  buffer space. This heuristic is made rigorous in the following proof that consists of an upper and lower bound. The proof of the upper bound is based on the fact that during the described events the only process that experiences losses is  $i$ . Equivalently, the total losses in the system are equal to the losses of process  $i$

$$\gamma_i^B(t) \approx (r_i + A_{-i} - c) \mathbf{1}\{\text{overflow due to } i\}.$$

The lower bound is conceptually easier, yet more tedious. The event of having flow  $i$  experiencing losses is intersected with events that process  $i$  has a long On period and all other processes exhibit its average behavior and have  $O(1)$  workloads.

**Proof. Upper bound:** Assume that all queueing processes below are equal to zero at time  $t = 0$ . Based on definition (2.3) of the instantaneous loss rate and its non-negativity, the following holds:

$$\begin{aligned} \gamma_i^B(t) &= (A_i(t) - c_i(t) - \dot{W}_i(t)) \mathbf{1}\{A_i(t) - c_i(t) > \dot{W}_i(t)\} \\ &\leq (r_i + A_{-i}(t) - c) \mathbf{1}\{A_i(t) - c_i(t) > \dot{W}_i(t)\}, \quad (4.1) \end{aligned}$$

where the last inequality follows from the fact that the buffer is full on event  $\{A_i(t) - c_i(t) > \dot{W}_i(t)\}$  and the instantaneous loss rate of a single flow is upper bounded by the total loss rate in the system  $A(t) - c$  that is further bounded by  $A(t) - c \leq r_i + A_{-i}(t) - c$ . Inequality (4.1), inclusion (2.4), Proposition 3.1 and the work-conserving property of the GPS scheduling scheme yield for all  $t \geq 0$

$$\begin{aligned} \gamma_i^B(t) &\leq (r_i + A_{-i}(t) - c) \mathbf{1}\left\{W_i(t) \geq B/N, \sum_{i=1}^N W_i(t) \geq B\right\} \\ &\leq (r_i + A_{-i}(t) - c) \mathbf{1}\left\{Q_{A_i}^{B,\phi_i c}(t) \geq B/N, Q_A^{B,c}(t) \geq B\right\}. \end{aligned}$$

Next, the preceding inequality and Proposition 3.2 result in

$$\begin{aligned} \gamma_i^B(t) &\leq (r_i + A_{-i}(t) - c) \mathbf{1}\left\{Q_{A_i}^{B,c-\rho_{-i}-\delta}(t) \geq \epsilon B\right\} \\ &\quad + (r - c) \mathbf{1}\left\{Q_{A_i}^{\infty,\phi_i c}(t) \geq B/N\right\} \\ &\quad \times \mathbf{1}\left\{Q_{A_{-i}}^{\infty,\rho_{-i}+\delta}(t) \geq (1 - \epsilon)B\right\}, \end{aligned}$$

where we select  $\epsilon \in (0, 1)$  and  $\delta \in (0, c - \rho)$ . Note that for any choice of  $\epsilon$  and  $\delta$  in the given intervals, all queueing processes in the last inequality are stable and converge in distribution to proper random variables as  $t \rightarrow \infty$ . Now, by independence of arrival processes, the last inequality renders

$$\begin{aligned} \Gamma_i^B &\leq (r_i + \rho_{-i} - c) \mathbb{P}\left[Q_{A_i}^{B,c-\rho_{-i}-\delta} \geq \epsilon B\right] \\ &\quad + (r - c) \mathbb{P}\left[Q_{A_i}^{\infty,\phi_i c} \geq B/N\right] \\ &\quad \times \mathbb{P}\left[Q_{A_{-i}}^{\infty,\rho_{-i}+\delta} \geq (1 - \epsilon)B\right]. \end{aligned}$$

Then, Theorem 3.1 and Lemma A.1 of Appendix A imply

$$\overline{\lim}_{B \rightarrow \infty} \frac{\Gamma_i^B}{A_{A_i}^{B,c-\rho_{-i}}} \leq \overline{\lim}_{B \rightarrow \infty} \frac{\mathbb{P}\left[Q_{A_i}^{B,c-\rho_{-i}-\delta} \geq \epsilon B\right]}{\mathbb{P}\left[Q_{A_i}^{B,c-\rho_{-i}} = B\right]}.$$

To complete the proof of the upper bound pass first  $\epsilon \uparrow 1$ , then  $\delta \downarrow 0$  and use Theorem 3.1.

*Lower bound:* Assume that all processes are in their stationary regimes unless otherwise specified. Recall that the loss rate is defined as  $\mathbb{E}\gamma_i^B(t)$ , where  $\gamma_i^B(t)$  is the instantaneous loss rate process of flow  $i$ . Clearly, for all  $0 < T < \infty$

$$\Gamma_i^B = \mathbb{E}\gamma_i^B(0) = T^{-1} \mathbb{E} \int_0^T \gamma_i^B(t) dt. \quad (4.2)$$

Next, for  $\epsilon > 0$  define quantity  $t_i \triangleq (1 + \epsilon)B / (r_i + \rho_{-i} - c)$  and two families of events

$$\zeta_i \triangleq \{A_i(0) = r_i, \inf\{t > 0: A_i(-t) = 0\} > t_i\},$$

$$\begin{aligned} \xi_i \triangleq & \bigcap_{j \neq i} \overline{\{A_j(0) = r_j, \inf\{t > 0: A_j(-t) = 0\} > B\}} \\ & \cap \{W_j(0) \leq B/(2N)\}. \end{aligned}$$

Event  $\zeta_i$  indicates that the  $i$ th flow is in On state at time  $t = 0$  and that the current On period lasted more than  $t_i$ . On event  $\{A_i(0) = r_i\}$ , the stationarity of arrival process  $A_i(t)$  implies  $\inf\{t > 0: A_i(-t) = 0\} \stackrel{d}{=} \tau_i^c$ . On the other hand,  $\xi_i$  indicates that all flows  $j \neq i$  do not have long On periods at time  $t = 0$  and the amount of their unfinished work is less than  $B/(2N)$ . Then, from (4.2), the loss rate for flow  $i$  can be lower bounded by

$$\Gamma_i^B \geq T^{-1} \mathbb{E} \left[ \int_0^T \gamma_i^B(t) dt \mathbf{1} \left\{ \zeta_i, \xi_i, \sum_{k=1}^N W_k(0) \geq B - K \right\} \right]. \quad (4.3)$$

Now, by using Proposition 3.3 and recalling that the considered system is work- and buffer-conserving we obtain

$$\begin{aligned} & \left\{ \zeta_i, \xi_i, K \geq B - \sum_{k=1}^N W_k(0) \right\} \\ &= \{ \zeta_i, \xi_i, K \geq B - Q_A^{B,c}(0) \} \\ &\geq \{ \zeta_i, \xi_i, K \geq Q_{c-r_i}^{\infty, A-i}(0) \}, \end{aligned} \quad (4.4)$$

where  $Q_{c-r_i}^{\infty, A-i}(t)$  denotes the workload of an infinite buffer queue with constant arrival rate  $c - r_i$ , service rate  $A_{-i}$  and the initial condition  $Q_{c-r_i}^{\infty, A-i}$

$(-t_i) = B$ . Next, for all  $B > 2NTr$  on event  $\xi_i$  the workloads of all flows other than  $i$  are smaller than  $B/N$  for all  $t \in (0, T)$  and, therefore, those flows do not experience loss of fluid. This leads to

$$\int_0^T \gamma_i(t) dt = \int_0^T \sum_{j=1}^N \gamma_j(t) dt \geq \int_0^T A(t) dt - cT - K,$$

where the last summand is due to the fact that at time  $t = 0$  the buffer can accommodate an additional  $K$  units of fluid (see (4.3)). Hence, the preceding inequality, (4.3) and (4.4) yield for all  $B > 2NTr$

$$\begin{aligned} \Gamma_i^B \geq & \mathbb{E} \left[ \left( -KT^{-1} + T^{-1} \int_0^T A(u) du - c \right) \right. \\ & \left. \times \mathbf{1} \{ Q_{c-r_i}^{\infty, A-i}(0) \leq K, \zeta_i, \xi_i \} \right]. \end{aligned}$$

Let  $t_i^*$  be the first time after  $t = 0$  that flow  $i$  is not generating fluid, i.e.  $t_i^* \triangleq \inf\{t > 0: A_i(t) = 0\}$ . Then, by the independence of flows

$$\begin{aligned} \Gamma_i^B \geq & -(KT^{-1} + c) \mathbb{P}[Q_{c-r_i}^{\infty, A-i}(0) \leq K, \xi_i] \mathbb{P}[\zeta_i] \\ & + T^{-1} \mathbb{E} \left[ \int_0^T A_{-i}(u) du \mathbf{1} \{ Q_{c-r_i}^{\infty, A-i}(0) \leq K, \xi_i \} \right] \mathbb{P}[\zeta_i] \\ & + r_i T^{-1} \mathbb{E} [ \min(T, t_i^*) \mathbf{1} \{ \zeta_i \} ] \mathbb{P}[Q_{c-r_i}^{\infty, A-i}(0) \leq K, \xi_i], \end{aligned} \quad (4.5)$$

where in the last term we used  $\int_0^T A_i(u) du \geq r_i \min(T, t_i^*)$  on event  $\zeta_i$ . Observe that since  $\tau_i^c \in \mathcal{I} \mathcal{R} \subset \mathcal{L}$  one obtains  $\mathbb{E}[\min(T, t_i^*) \mathbf{1} \{ \zeta_i \}] \sim T \mathbb{P}[\zeta_i]$  as  $B \rightarrow \infty$ . This, in conjunction with the independence of the arrival processes and the fact that  $\lim_{B \rightarrow \infty} \mathbb{P}[\xi_i] = 1$ , implies

$$\begin{aligned} \underline{\lim}_{B \rightarrow \infty} \frac{\Gamma_i^B}{\mathbb{P}[\zeta_i]} \geq & (-KT^{-1} + r_i - c + \rho_{-i}) \\ & \times \underline{\lim}_{B \rightarrow \infty} \mathbb{P}[Q_{c-r_i}^{\infty, A-i}(0) \leq K] \\ & - r_i \overline{\lim}_{B \rightarrow \infty} \mathbb{P}[Q_{c-r_i}^{\infty, A-i}(0) > K]. \end{aligned}$$

The preceding inequality holds for all  $T > 0$  and, therefore, passing  $T \rightarrow \infty$  and using Theorem 3.1 render

$$\begin{aligned}
& \lim_{B \rightarrow \infty} \frac{\Gamma_i^B}{A_{A_i}^{B, c - \rho_{-i}}} \\
& \geq \left( 1 - \frac{2r}{r_i + \rho_{-i} - c} \overline{\lim}_{B \rightarrow \infty} \mathbb{P}[Q_{c-r_i}^{\infty, A_{-i}}(0) > K] \right) \\
& \quad \times \lim_{B \rightarrow \infty} \frac{\mathbb{P}\left[\tau_i^c > \frac{(1+\epsilon)B}{r_i + \rho_{-i} - c}\right]}{\mathbb{P}\left[\tau_i^c > \frac{B}{r_i + \rho_{-i} - c}\right]}. \quad (4.6)
\end{aligned}$$

Next, the standard queueing reflection mapping argument provides a representation for

$$\begin{aligned}
Q_{c-r_i}^{\infty, A_{-i}}(0) = \max \left( \sup_{-t_i \leq s \leq 0} \left\{ |s|(c - r_i) - \int_s^0 A_{-i} du \right\}, \right. \\
\left. B + (c - r_i)t_i - \int_{-t_i}^0 A_{-i} du \right)
\end{aligned}$$

and, hence, by the Strong Law of Large numbers and  $c - r_i < \rho_{-i}$  one obtains for all  $\epsilon > 0$

$$\lim_{K \rightarrow \infty} \overline{\lim}_{B \rightarrow \infty} \mathbb{P}[Q_{c-r_i}^{\infty, A_{-i}}(0) > K] = 0.$$

Finally, by setting  $\epsilon \downarrow 0$  and then  $K \rightarrow \infty$  in (4.6) the lower bound follows by the preceding limit. This completes the proof of the theorem.  $\square$

At this point we would like to discuss possible generalizations of the preceding theorem. The strict stability condition  $c\phi_i > \rho_i$  for all  $i$  represents a natural engineering condition. However, from a theoretical perspective the behavior of the system remains unclear if one or more flows have higher average demands than their minimum guaranteed rates. The following corollary represents an easy extension of this type. It states that a flow with guaranteed service rate lower than its expected rate will not be asymptotically affected by other flows if the tail of that flow is sufficiently heavy. In general, the problems of this kind are difficult and remain open.

**Corollary 4.1.** *Let  $\tau_j^c \in \mathcal{I}\mathcal{R}$  for  $1 \leq j \leq N$  and  $r_i + \rho_{-i} > c > \rho$  for some flow  $i$ . If  $\phi_j c > \rho_j$  for all  $j \neq i$  and  $\prod_{j \in \mathcal{Q}} \mathbb{P}[\tau_j^c > x] = o(\mathbb{P}[\tau_i^c > x])$  for all sets  $\mathcal{Q} \subseteq \{1, \dots, N\}$  that satisfy  $\sum_{j \in \mathcal{Q}} r_j + \sum_{j \notin \mathcal{Q}} \rho_j \geq c$ , then as  $B \rightarrow \infty$*

$$\Gamma_i^B \sim A_{A_i}^{B, c - \rho_{-i}}.$$

**Proof.** The upper bound is a direct consequence of inclusion (2.4) and Theorem 3.2. The proof of the lower bound is the same as in Theorem 4.1.  $\square$

Careful examination of the proof of Theorem 4.1 shows that it holds for much more general buffer management policies. In particular, this includes the GPS-like rules for buffer management, which can be utilized to improve the behavior of the system for small queue sizes. More generally, Theorem 4.1 appears to hold for a large class of buffer-conserving policies under which the guaranteed buffer space for each process is an increasing function of the total buffer size.

It is worth mentioning that when condition  $r_i + \rho_{-i} > c$  fails, the most likely overflow for process  $i$  may be caused jointly by flow  $i$  and a subset of other flows. Unfortunately, it appears that the expression for the loss rate in the general case depends on the buffer management policy as well as the parameters of other processes. Hence, the answers and analysis under these assumptions are expected to be more involved.

## 5. Numerical examples

In this section we demonstrate that our analytical formulas match well the simulation experiments. Consider a fluid queue with capacity  $c = 2.5$  and finite buffer  $B$  shared by five On–Off flows. Since the asymptotic results are insensitive to the distribution of Off periods, we choose their distribution to be exponential, i.e.,  $\mathbb{P}[\eta > x] = e^{-\mu x}$ ,  $x \geq 0$ . On periods are selected from the Pareto family,  $\mathbb{P}[\tau > x] = x^{-\alpha}$ ,  $x \geq 1$ ,  $\alpha > 1$ . The peak rates of On–Off flows are defined by vector  $(4, 2, 2, 3, 1)$  and the On probabilities are chosen to be  $p_i = 0.1$  for all five flows. The length of the simulated sample path in both examples is set to  $10^9$ .

**Example 5.1.** Let the distributions of On periods be defined by  $\alpha = (1.6, 1.5, 1.6, 1.5, 1.6)$ . The work of the GPS mechanism is fully determined by the weight vector  $\phi = (0.3, 0.1, 0.3, 0.2, 0.1)$ . Clearly, the conditions of Theorem 4.1 are satisfied for the first flow. For this flow we simulated the loss rates



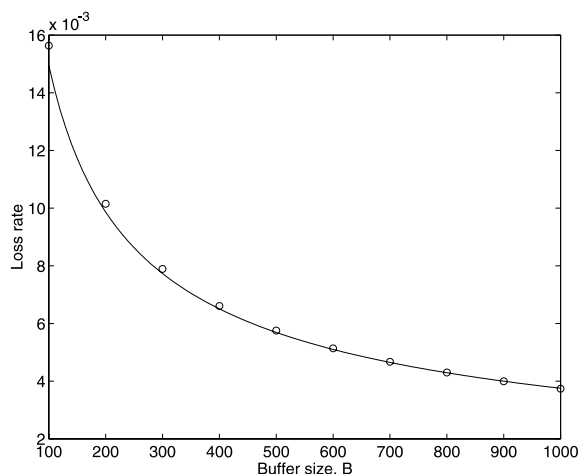


Fig. 3. Illustration for Example 5.1.

for buffer sizes  $B = 100, 200, \dots, 1000$ . The results of the simulation are presented in Fig. 3 with “o” symbols. The approximation of the loss rate is plotted in the same figure with a solid line.

**Example 5.2.** Here, consider the system from the previous example with  $\alpha = (1.6, 3.0, 3.0, 3.4, 1.9)$  and  $\phi = (0.1, 0.2, 0.3, 0.3, 0.1)$ . In this example it is easy to verify that the conditions of Corollary 4.1 are satisfied for the first flow. The simulation results (“o” symbols) and the approximation (solid line) for the loss rate are plotted on Fig. 4.

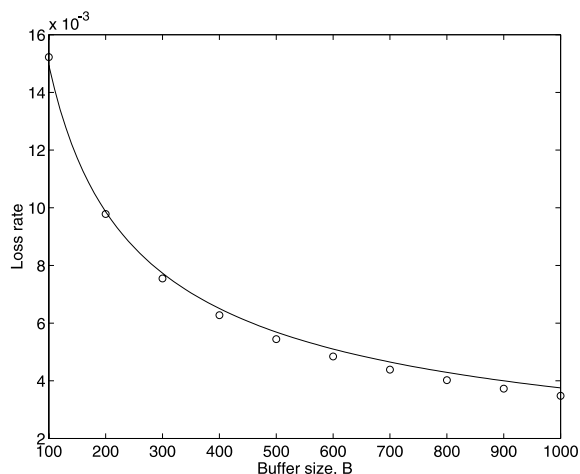


Fig. 4. Illustration for Example 5.2.

It is apparent from both figures that the derived approximations are in agreement with the simulated results. Furthermore, we would like to point out that the asymptotic formulas are quite precise even though the estimated probabilities are relatively high, in the range of  $10^{-2}$ – $10^{-3}$ .

## 6. Concluding remarks

In this paper we analyzed a system of  $N$  independent heavy-tailed On–Off flows that share a server and a common buffer. The server capacity is divided according to the GPS scheduling discipline and the buffer sharing is unrestricted, unless the buffer is full. When the buffer is full, we introduce a low complexity workload management policy that discards the necessary amount of fluid from the most demanding flows. Each flow is assumed to have its minimum rate guarantee larger than its long-time average rate. Our main result shows that asymptotically each flow experiences the same loss rate as if it were served in isolation with constant capacity and the whole buffer space  $B$  exclusively dedicated to it.

The main novelty of our result is that the analyzed system behaves as if it had  $N$  times larger buffer than it actually did. The result indicates that the buffer sharing can be very beneficial in the case of heavy-tailed traffic. The new insight provides an additional guideline for deciding on the buffer distribution between the access and core network switching elements. Furthermore, we observe that the result raises an interesting problem of socially fair buffer pricing. Everybody gets a full share of the resource, but it is not quite clear who and how much one should pay for it.

The analyzed system represents a baseline model of providing GPS-based, such as weighted fair queueing, differentiated services using an open loop control mechanism, e.g. UDP. The analysis of our model with the addition of TCP-like closed loop flow control appears at the moment out of reach. However, even in the closed loop *peak rate limited systems*, the time and behavior of the buffer content until an overflow occurs is very similar to the one using an open loop control. The two systems will behave differently during the overflow period and

then, assuming stability, both buffers will drain to zero. Then, until the next overflow, the systems will behave almost the same, etc. Thus, one can expect the same number of overflow episodes in both systems, but with a different dynamics of losses. Hence, one is tempted to conjecture, that the loss probabilities in a system similar to ours with a TCP-like flow control, will be of the same order as in Theorem 4.1, but with a different constant of proportionality. Therefore, one may expect analogous multiplexing gains of having larger buffers in core routers that are carrying peak rate limited closed loop controlled traffic. However, the positive economics of the buffer multiplexing gains has to be balanced with a potential increase in cost for high speed buffers that are necessary for core routers, and, possibly other switch design tradeoffs.

## Appendix A. Heavy-tailed distributions

**Definition A.1.** A non-negative random variable  $X$  is called long-tailed,  $X \in \mathcal{L}$ , if  $\mathbb{P}[X > x - y] \sim \mathbb{P}[X > x]$  as  $x \rightarrow \infty$ ,  $\forall y \in \mathbb{R}$ .

**Definition A.2.** A non-negative random variable  $X$  is called subexponential,  $X \in \mathcal{S} \subset \mathcal{L}$ , if  $\mathbb{P}[X + Y > x] \sim \mathbb{P}[X > x]$  as  $x \rightarrow \infty$ , where  $Y$  is an independent copy of  $X$ .

**Definition A.3.** A non-negative random variable  $X$  is called intermediately regularly varying,  $X \in \mathcal{IR} \subset \mathcal{S} \subset \mathcal{L}$ , if

$$\lim_{\eta \downarrow 1} \overline{\lim}_{x \rightarrow \infty} \frac{\mathbb{P}[X > \eta x]}{\mathbb{P}[X > x]} = 1.$$

**Lemma A.1.** Let  $X \in \mathcal{IR}$ ,  $\eta \in (0, 1)$ , then

$$\sup_{x \in (0, \infty)} \frac{\mathbb{P}[X > \eta x]}{\mathbb{P}[X > x]} < \infty.$$

**Proof.** Follows immediately from the definition.  $\square$

## References

- [1] D. Bertsimas, I.C. Paschalidis, J.N. Tsitsiklis, Large deviations analysis of the general processor sharing policy, *Queueing Systems* 32 (1999) 319–349.
- [2] S. Borst, O. Boxma, P. Jelenković, Generalized processor sharing with long-tailed traffic sources, in: P. Key, D. Smith (Eds.), *Teletraffic Engineering in a Competitive World*, Proc. ITC-16, Edinburgh, UK, North-Holland, Amsterdam, 1999, pp. 345–354.
- [3] S. Borst, O. Boxma, P. Jelenković, Induced burstiness in generalized processor sharing queues with long-tailed traffic flows, in: 37th Annual Allerton Conference on Communication, Control, and Computing, 1999.
- [4] S. Borst, O. Boxma, P. Jelenković, Asymptotic behavior of generalized processor sharing with long-tailed traffic sources, in: *INFOCOM 2000*, Tel-Aviv, Israel, 2000.
- [5] S. Borst, O. Boxma, P. Jelenković, Coupled processors with regularly varying service times, in: *INFOCOM 2000*, Tel-Aviv, Israel, 2000.
- [6] S. Borst, O. Boxma, P. Jelenković, Reduced-load equivalence and induced burstiness in GPS queues with long-tailed traffic flows, *Queueing Systems*, to appear.
- [7] M. Crovella, A. Bestavros, Self-similarity in world wide web traffic: evidence and possible causes, *IEEE/ACM Transactions on Networking* 5 (6) (1997) 835–846.
- [8] P. Dupuis, K. Ramanan, A Skorohod problem formulation and large deviation analysis of a processor sharing model, *Queueing Systems* 28 (1–3) (1998) 109–124.
- [9] G. Fayolle, I. Mitrani, R. Iasnogorodski, Sharing a processor among many job classes, *Journal of ACM* 27 (3) (1980) 519–532.
- [10] D.P. Heyman, T.V. Lakshman, Source models for VBR broadcast-video traffic, *IEEE/ACM Transactions on Networking* 4 (1996) 40–48.
- [11] P. Jelenković, P. Momčilović, Asymptotic loss probability in a finite buffer fluid queue with heterogeneous heavy-tailed on-off processes, *Annals of Applied Probability*, in press.
- [12] P. Jelenković, P. Momčilović, Capacity regions for network multiplexers with heavy-tailed fluid on-off sources, in: *INFOCOM 2001*, Anchorage, Alaska, April 2001.
- [13] P.R. Jelenković, Subexponential loss rates in a GI/GI/1 queue with applications, *Queueing Systems* 33 (1999) 91–123.
- [14] P.R. Jelenković, A.A. Lazar, Asymptotic results for multiplexing subexponential on-off processes, *Advances Applied Probability* 31 (2) (1999).
- [15] P.R. Jelenković, A.A. Lazar, N. Semret, The effect of multiple time scales and subexponentiality of MPEG video streams on queueing behavior, *IEEE JSAC* 15 (6) (1997) 1052–1071.
- [16] W.E. Leland, M.S. Taqqu, W. Willinger, D.V. Wilson, On the self-similar nature of Ethernet traffic, in: *SIGCOMM'93*, 1993, pp. 183–193.
- [17] R.M. Loynes, The stability of a queue with non-independent inter-arrival and service times, *Proceedings of the Cambridge Philosophical Society* 58 (1962) 497–520.
- [18] L. Massoulié, Large deviations estimates for polling and weighted fair queueing service systems, *Advances in Performance Analysis* 2 (2) (1999) 103–128.
- [19] K. Park, W. Willinger (Eds.), *Self-Similar Network Traffic and Performance Evaluation*, John Wiley and Sons, 2000.

- [20] Z.-L. Zhang, Large deviations and the generalized processor sharing scheduling for a multiple-queue system, *Queueing Systems* 28 (1998) 349–376.
- [21] Z.-L. Zhang, D. Towsley, J. Kurose, Statistical analysis of the generalized processor sharing scheduling discipline, *IEEE JSAC* 13 (6) (1995).

**Predrag R. Jelenković** was born in Kraljevo, Serbia, in 1966. He received the B.S. degree from the University of Belgrade, Serbia, in 1991, and the M.S., M.Ph., and Ph.D. degrees from Columbia University, New York, in 1993, 1995, and 1996, respectively, all in Electrical Engineering.

Immediately after his graduation in 1996 he joined the Mathematics of Networks and Systems Department, at Bell Labs, Lucent Technologies, where he worked for two years. In

1998 he became the faculty of the Department of Electrical Engineering at Columbia University, where he is currently an Associate Professor.

Dr. Jelenković was a recipient of the Best Student Paper award at the International Teletraffic Congress in 1997 and the NSF Presidential Early Career Award for Scientists and Engineers in 2000. His research focuses on mathematical foundations of communication networks under emerging traffic paradigms. <http://www.comet.columbia.edu/~predrag>.

**Petar Momčilović** received the B.S. degree from the Moscow Power Engineering Institute, Russia, in 1997, and the M.S. degree with distinction from the Department of Electrical Engineering at Columbia University, New York, in 2000. Currently, he is a Ph.D. candidate in the same department. His research interests include mathematical modeling, analysis and control of computer communications networks and systems. <http://www.comet.columbia.edu/~petar>.