Generalized Processor Sharing with Long-Tailed Traffic Sources ¶

Sem Borst[*,†,‡], Onno Boxma[†,‡], Predrag Jelenković[**]

[*]Bell Laboratories, Lucent Technologies, P.O. Box 636, Murray Hill, NJ 07974, USA

[**]Dept. of Electrical Engineering, Columbia University, New York, NY 10027, USA

[†]CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

[‡]Department of Mathematics & Computing Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

We analyze the queueing behavior of long-tailed traffic sources under the Generalized Processor Sharing (GPS) discipline. GPS-based scheduling algorithms, such as Weighted Fair Queueing, have emerged as important mechanisms for accommodating heterogeneous quality-of-service requirements in integrated-services networks.
Under mild assumptions, we show that the tail behavior of the buffer content of an individual source with long-tailed traffic characteristics is equivalent to the tail behavior when that source is served in isolation at a *constant* rate which is equal to the link rate minus the aggregate average rate of all other sources. Thus, asymptotically, the buffer content of the source is only affected by the traffic characteristics of the other sources through their *aggregate average* rate. In particular, the source is essentially immune to excessive activity of sources with 'heavier'-tailed traffic characteristics. This suggests that GPS-based scheduling algorithms provide an effective mechanism for extracting high multiplexing gains, while protecting individual connections.

1. Introduction

An increasing body of literature on statistical data analysis has provided evidence of long-tailed (subexponential) traffic characteristics in high-speed communication networks. Early discoveries on the long-range dependence of Ethernet traffic, attributed to long-tailed file-size distributions, were reported in [14]. Long-tailed characteristics of the scene length distribution of MPEG video streams were explored in [13].

These empirical findings have encouraged theoretical developments in the modeling and queueing analysis of long-tailed traffic phenomena. We refer to Boxma & Dumas [7] for a comprehensive survey on fluid queues with long-tailed arrival processes. See also Jelenković [10] for an extensive list of references on subexponential queueing models.

¶ITC-16, Edinburgh, UK, June 7-11, 1999.

Despite significant progress, however, the practical implications are not yet thoroughly understood, in particular issues relating to control and priority mechanisms in the network. To gain a better understanding of those issues, the present paper analyzes the queueing behavior of long-tailed traffic sources under the Generalized Processor Sharing (GPS) discipline. GPS serves as a design paradigm for packet scheduling algorithms, such as Weighted Fair Queueing, see for instance Parekh & Gallager [17].

The impact of priority mechanisms on long-tailed traffic phenomena has received relatively little attention. Some recent studies have investigated the effect of priorities on the waiting-time distribution in the classical M/G/1 queue. For FCFS, it is well-known [8] that the waiting-time tail is regularly varying of index $1-\nu$ iff the service time tail is regularly varying of index $-\nu$. For Processor Sharing, the waiting-time tail turns out to be regularly varying of the *same* index as the service time tail [20]. In the case of Processor Sharing with several customer classes, Zwart [19] recently showed that the sojourn time distribution of a class-i customer is regularly varying of index $-\nu_i$ iff the service time distribution of that class is regularly varying of index $-\nu_i$, *regardless* of the service time distributions of other customer classes.

In the present paper, we consider the Generalized Processor Sharing (GPS) discipline. GPS-based scheduling algorithms, such as Weighted Fair Queueing, have emerged as an important mechanism for accommodating heterogeneous quality-of-service requirements in integrated-services networks. The queueing analysis of GPS is extremely difficult. Interesting partial results for exponential traffic models were obtained in Bertsimas *et al.* [4], Dupuis & Ramanan [9], Massoulié [15], and Zhang *et al.* [18].

Here, we focus on non-exponential traffic models. We show that, under mild assumptions, the tail behavior of the buffer content of an individual source with long-tailed traffic characteristics is equivalent to the tail behavior when that source is served in isolation at a constant rate which is equal to the link rate minus the aggregate average rate of all other sources. Thus, asymptotically, the buffer content of the source is only influenced by the traffic characteristics of the other sources through their *aggregate average* rate. In particular, the source is largely insensitive to extreme activity patterns of sources with 'heavier'-tailed traffic characteristics. This suggests that GPS-based scheduling algorithms provide an effective mechanism for extracting high multiplexing gains, while protecting individual connections.

The isolation of individual connections in GPS is somewhat similar to that for regular processor sharing observed in [19] mentioned above, and should be contrasted with an ordinary M/G/1 queue with N arrival streams served on a FCFS basis. In the latter scenario, if the service time distributions are regularly varying, then the tail of the workload and waiting-time distribution is determined by the *heaviest* of the tails of those distributions [8]. This remains true when the N arrival streams are served according to a non-preemptive priority rule [1,6]. Under preemptive resume, a particular class is completely oblivious of the presence of lower-priority classes, but is not protected from the influence of heavy-tailed characteristics of higher-priority classes.

The remainder of the paper is organized as follows. In Section 2, we present a detailed model description. In Section 3, we derive general lower and upper bounds on the buffer content distribution of an individual source. We then show, in Section 4, that for long-tailed source characteristics, the lower and upper bounds have the same asymptotic behavior, yielding an explicit characterization of the tail of the buffer content distribution.

2. Model description

Consider N sources sharing a link of unit rate. Traffic from the sources is served in accordance with the Generalized Processor Sharing (GPS) discipline, which operates as follows. There are weights $\phi_1, \ldots, \phi_N$ associated with each of the sources, with $\sum_{i=1}^{N} \phi_i = 1$. Denote by $V_i(t)$ the buffer content of source i at time t. Let $\mathbf{V}_i$ be a stochastic variable with the limiting distribution of $V_i(t)$ for $t \to \infty$. If all the sources are backlogged at time t, i.e., $V_i(t) > 0$ for all $i = 1, \ldots, N$, then source i is served at rate ϕ_i. If some of the sources are not backlogged, however, then the excess capacity is redistributed among the backlogged sources in accordance with their respective weights. We refer to Dupuis & Ramanan [9] for a formal description of the evolution of the buffer content process. Notice that the GPS discipline has the following two properties: (i) it is work-conserving, i.e., it serves at the full link rate whenever any of the sources is backlogged; (ii) it guarantees minimum rates $\phi_1, \ldots, \phi_N$, i.e., it serves source i at least at rate ϕ_i whenever source i is backlogged.

Denote by $A_i(s,t)$ the amount of traffic generated by source i during the time interval $[s,t]$, and let $A(s,t) := \sum_{i=1}^{N} A_i(s,t)$ be the total amount of traffic generated during that time interval. Denote by ρ_i the traffic intensity of source i as will be defined below in detail for the two traffic models that we consider. Define $\rho := \sum_{i=1}^{N} \rho_i$ as the total traffic intensity. Define $V(t) := \sum_{i=1}^{N} V_i(t)$ as the aggregate buffer content at time t. From property (i),

$$V(t) = \sup_{0 \le s \le t} \{A(s,t) - (t-s)\} \qquad \text{for all} \quad t \ge 0. \tag{1}$$

For any $c \ge 0$, denote by $W_i^c(t) := \sup_{0 \le s \le t}\{A_i(s,t) - c(t-s)\}$ the buffer content of source i at time t if it were served in isolation at rate c. For $c > \rho_i$, let $\mathbf{W}_i^c$ be a stochastic variable with the limiting distribution of $W_i^c(t)$ for $t \to \infty$. From property (ii),

$$V_i(t) \le W_i^{\phi_i}(t) \qquad \text{for all} \quad t \ge 0. \tag{2}$$

Throughout the paper, we make the following assumption.

Assumption 2.1 *The traffic intensities and weights satisfy $\rho_i < \phi_i$ for all $i = 1, \ldots, N$.*

Notice that the above assumption implies $\rho < 1$ which ensures stability of the sources. For a formal stability proof, we refer to Dupuis & Ramanan [9].

Before describing the traffic model, we first introduce some further notation. Throughout the paper, for any two real functions $g(\cdot)$ and $h(\cdot)$, we use the customary notation $g(x) \sim h(x)$ as $x \to \infty$ to denote $\lim_{x\to\infty} g(x)/h(x) = 1$, or equivalently, $g(x) = h(x)(1 + o(1))$ as $x \to \infty$. For any stochastic variable $\mathbf{X}$ with distribution function $F(\cdot)$, $\mathrm{E}\mathbf{X} < \infty$, denote by $F^r(\cdot)$ the distribution function of the residual lifetime of $\mathbf{X}$, i.e., $F^r(x) = \frac{1}{\mathrm{E}\mathbf{X}} \int_0^x (1 - F(y))\mathrm{d}y$.

The classes of *long-tailed* and *subexponential* distributions are denoted with the symbols $\mathcal{L}$ and $\mathcal{S}$, respectively. The definitions of these classes are given in the appendix.

We now describe the two traffic models that we consider.

2.1. Instantaneous arrivals

Here, sources generate instantaneous traffic bursts according to independent renewal processes. The interarrival times between bursts of source i are generally distributed with mean $1/\lambda_i$. The burst sizes of source i have distribution $B_i(\cdot)$ with mean β_i. Thus, the traffic intensity of source i is $\rho_i = \lambda_i \beta_i$. Let $\mathbf{B}_i^r$ be a stochastic variable with distribution $B_i^r(\cdot)$.

The next result is immediate from [16].

Theorem 2.1 *If $B_i^r(\cdot) \in \mathcal{S}$ and $\rho_i < c$, then*

$$\Pr\{\mathbf{W}_i^c > x\} \sim \frac{\rho_i}{c - \rho_i} \Pr\{\mathbf{B}_i^r > x\} \qquad \text{as} \quad x \to \infty.$$

2.2. On-Off processes

Here, sources generate traffic according to independent On-Off processes, alternating between On- and Off-periods. The Off-periods of source i are generally distributed with mean $1/\lambda_i$. The On-periods of source i have distribution $A_i(\cdot)$ with mean α_i. While On, source i produces traffic at a constant rate r_i, so the mean burst size is $\alpha_i r_i$. The fraction of time that source i is Off is $p_i = 1/(1 + \lambda_i \alpha_i)$. The traffic intensity of source i is $\rho_i = \lambda_i \alpha_i r_i / (1 + \lambda_i \alpha_i)$. Let $\mathbf{A}_i^r$ be a stochastic variable with distribution $A_i^r(\cdot)$.

The next result is established in [11].

Theorem 2.2 *If $A_i^r(\cdot) \in \mathcal{S}$ and $r_i > c$, then*

$$\Pr\{\mathbf{W}_i^c > x\} \sim p_i \frac{\rho_i}{c - \rho_i} \Pr\{\mathbf{A}_i^r > x/(r_i - c)\} \qquad \text{as} \quad x \to \infty.$$

3. Bounds

We now derive general bounds for the buffer content distribution, which we will use in the next section to analyze the tail behavior. We first introduce some further notation. For notational convenience, define $A_{-i}(s,t) := A(s,t) - A_i(s,t) = \sum_{j \neq i} A_j(s,t)$ as the amount of traffic generated by all the other sources during the time interval $[s,t]$. Also, denote by $\rho_{-i} = \rho - \rho_i = \sum_{j \neq i} \rho_j$ the aggregate average rate of all the other sources. For any c, define $U_{-i}^c(t) := \sup_{0 \leq s \leq t} \{c(t-s) - A_{-i}(s,t)\}$. For $c < \rho_{-i}$, let $\mathbf{U}_{-i}^c$ be a stochastic variable with the limiting distribution of $U_{-i}^c(t)$ as $t \to \infty$.

We first present a lower bound for the buffer content distribution.

Lemma 3.1 *(Lower bound) For any $\delta > 0$ and y,*

$$\Pr\{\mathbf{V}_i > x\} \geq \Pr\{\mathbf{W}_i^{1-\rho_{-i}+\delta} > x + y\} \Pr\{\mathbf{U}_{-i}^{\rho_{-i}-\delta} + \sum_{j \neq i} \mathbf{W}_j^{\phi_j} < y\}. \tag{3}$$

Proof Using equations (1), (2) we obtain, for any $\theta \leq 1$,

$$\begin{aligned}
V_i(t) &= V(t) - \sum_{j \neq i} V_j(t) \\
&\geq \sup_{0 \leq s \leq t} \{A(s,t) - (t-s)\} - \sum_{j \neq i} W_j^{\phi_j}(t) \\
&= \sup_{0 \leq s \leq t} \{A_i(s,t) - (1-\theta)(t-s) + A_{-i}(s,t) - \theta(t-s)\} - \sum_{j \neq i} W_j^{\phi_j}(t) \\
&\geq \sup_{0 \leq s \leq t} \{A_i(s,t) - (1-\theta)(t-s)\} + \inf_{0 \leq s \leq t} \{A_{-i}(s,t) - \theta(t-s)\} - \sum_{j \neq i} W_j^{\phi_j}(t) \\
&= W_i^{1-\theta}(t) - U_{-i}^{\theta}(t) - \sum_{j \neq i} W_j^{\phi_j}(t) \qquad \text{for all} \quad t \geq 0.
\end{aligned}$$

In particular, taking $\theta = \rho_{-i} - \delta$, we have

$$V_i(t) \geq W_i^{1-\rho_{-i}+\delta}(t) - U_{-i}^{\rho_{-i}-\delta}(t) - \sum_{j \neq i} W_j^{\phi_j}(t) \qquad \text{for all} \quad t \geq 0.$$

Thus, in the stationary regime, using independence,

$$\begin{aligned}
\Pr\{\mathbf{V}_i > x\} &\geq \Pr\{\mathbf{W}_i^{1-\rho_{-i}+\delta} - \mathbf{U}_{-i}^{\rho_{-i}-\delta} - \sum_{j \neq i} \mathbf{W}_j^{\phi_j} > x\} \\
&\geq \Pr\{\mathbf{W}_i^{1-\rho_{-i}+\delta} > x+y, \mathbf{U}_{-i}^{\rho_{-i}-\delta} + \sum_{j \neq i} \mathbf{W}_j^{\phi_j} < y\} \\
&= \Pr\{\mathbf{W}_i^{1-\rho_{-i}+\delta} > x+y\} \Pr\{\mathbf{U}_{-i}^{\rho_{-i}-\delta} + \sum_{j \neq i} \mathbf{W}_j^{\phi_j} < y\}.
\end{aligned}$$

□

We now provide an upper bound for the buffer content distribution. For any $c \geq 0$, define $W_{-i}^c(t) := \sup_{0 \leq s \leq t} \{A_{-i}(s,t) - c(t-s)\}$ as the aggregate buffer content of all sources other than i if they were served at rate c in a work-conserving manner. For $c > \rho_{-i}$, let $\mathbf{W}_{-i}^c$ be a stochastic variable with the limiting distribution of $W_{-i}^c(t)$ as $t \to \infty$.

Lemma 3.2 *(Upper bound) For any $\delta > 0$ and y, with $\delta < 1 - \rho$,*

$$\Pr\{\mathbf{V}_i > x\} \leq \Pr\{\mathbf{W}_i^{1-\rho_{-i}-\delta} > x - y\} + \Pr\{\mathbf{W}_i^{\phi_i} > x\} \Pr\{\mathbf{W}_{-i}^{\rho_{-i}+\delta} > y\}. \tag{4}$$

Proof *Upper bound.* For any $\theta \leq 1$, we have, using equation (1),

$$\begin{aligned}
V_i(t) &\leq V(t) \\
&= \sup_{0 \leq s \leq t} \{A(s,t) - (t-s)\} \\
&= \sup_{0 \leq s \leq t} \{A_i(s,t) - (1-\theta)(t-s) + A_{-i}(s,t) - \theta(t-s)\} \\
&\leq \sup_{0 \leq s \leq t} \{A_i(s,t) - (1-\theta)(t-s)\} + \sup_{0 \leq s \leq t} \{A_{-i}(s,t) - \theta(t-s)\} \\
&= W_i^{1-\theta}(t) + W_{-i}^{\theta}(t) \qquad \text{for all} \quad t \geq 0.
\end{aligned}$$

Invoking (2), and setting $\theta = \rho_{-i} + \delta$, we obtain

$$V_i(t) \leq \min\{W_i^{\phi_i}(t), W_i^{1-\rho_{-i}-\delta}(t) + W_{-i}^{\rho_{-i}+\delta}(t)\} \qquad \text{for all} \quad t \geq 0.$$

Thus, in the stationary regime, using independence,

$$\begin{aligned}
\Pr\{\mathbf{V}_i > x\} &\leq \Pr\{\mathbf{W}_i^{\phi_i} > x, \mathbf{W}_i^{1-\rho_{-i}-\delta} + \mathbf{W}_{-i}^{\rho_{-i}+\delta} > x\} \\
&\leq \Pr\{\mathbf{W}_i^{\phi_i} > x, \mathbf{W}_i^{1-\rho_{-i}-\delta} > x - y \text{ or } \mathbf{W}_{-i}^{\rho_{-i}+\delta} > y\} \\
&\leq \Pr\{\mathbf{W}_i^{1-\rho_{-i}-\delta} > x - y\} + \Pr\{\mathbf{W}_i^{\phi_i} > x, \mathbf{W}_{-i}^{\rho_{-i}+\delta} > y\} \\
&= \Pr\{\mathbf{W}_i^{1-\rho_{-i}-\delta} > x - y\} + \Pr\{\mathbf{W}_i^{\phi_i} > x\}\Pr\{\mathbf{W}_{-i}^{\rho_{-i}+\delta} > y\}.
\end{aligned}$$

□

4. Asymptotic behavior

We now use the bounds from the previous section to determine the tail behavior of the buffer content distribution under long-tailed traffic assumptions.

4.1. Instantaneous arrivals

We first characterize the tail distribution of the buffer content for sources with instantaneous subexponential arrivals.

Theorem 4.1 *If $B_i^r(\cdot) \in \mathcal{S}$, then*

$$\Pr\{\mathbf{V}_i > x\} \sim \Pr\{\mathbf{W}_i^{c_i} > x\} \qquad \text{as} \quad x \to \infty,$$

with $c_i = 1 - \rho_{-i}$ representing the link rate minus the aggregate average rate of all other sources.

Theorem 4.1 says that the tail behavior of the buffer content of an individual source i (with subexponential burst size distribution) is equivalent to the tail behavior when that source is served in isolation at a *constant* rate which is equal to the link rate minus the aggregate average rate of all other sources. Thus, asymptotically, the buffer content of source i is only affected by the burst size distribution of the other sources through their *aggregate average* rate. This is reminiscent of the result obtained in [11] for multiplexing exponential with subexponential processes and the 'reduced-load equivalence' established by Agrawal *et al.* [2]. However, for the latter results, the authors needed to impose bounding conditions on the variability of other sources ($\neq i$). Here, due to the nature of GPS scheduling, such conditions are not required. In particular, source i is essentially immune to excessive activity of other sources, even when they have 'heavier' burst size distributions! The exact values of the GPS weights are not relevant either, as long as they are larger than the traffic intensities (Assumption 2.1).

Remark 4.1 Theorem 4.1 can be generalized to Markov-modulated arrivals using an extension to Theorem 2.1 in [12] or restated for even more general arrival processes [3].

Proof of Theorem 4.1 The lower bound in Lemma 3.1 implies, for any $\delta > 0$ and y,

$$\frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{B}_i^r > x\}} \geq \frac{\Pr\{\mathbf{W}_i^{1-\rho_{-i}+\delta} > x+y\}}{\Pr\{\mathbf{B}_i^r > x+y\}} \frac{\Pr\{\mathbf{B}_i^r > x+y\}}{\Pr\{\mathbf{B}_i^r > x\}} \Pr\{\mathbf{U}_{-i}^{\rho_{-i}-\delta} + \sum_{j\neq i} \mathbf{W}_j^{\phi_j} < y\}.$$

Using Definition 1 and Theorem 2.1,

$$\liminf_{x\to\infty} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{B}_i^r > x\}} \geq \frac{\rho_i}{1-\rho+\delta} \Pr\{\mathbf{U}_{-i}^{\rho_{-i}-\delta} + \sum_{j\neq i} \mathbf{W}_j^{\phi_j} < y\}.$$

Letting $y \to \infty$, and then $\delta \downarrow 0$, we have

$$\liminf_{x\to\infty} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{B}_i^r > x\}} \geq \frac{\rho_i}{1-\rho}. \tag{5}$$

Similarly, the upper bound in Lemma 3.2 gives, for any $\delta > 0$ and y, with $\delta < 1-\rho$,

$$\frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{B}_i^r > x\}} \leq \frac{\Pr\{\mathbf{W}_i^{1-\rho_{-i}-\delta} > x-y\}}{\Pr\{\mathbf{B}_i^r > x-y\}} \frac{\Pr\{\mathbf{B}_i^r > x-y\}}{\Pr\{\mathbf{B}_i^r > x\}} + \frac{\Pr\{\mathbf{W}_i^{\phi_i} > x\}}{\Pr\{\mathbf{B}_i^r > x\}} \Pr\{\mathbf{W}_{-i}^{\rho_{-i}+\delta} > y\},$$

so that

$$\limsup_{x\to\infty} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{B}_i^r > x\}} \leq \frac{\rho_i}{1-\rho-\delta} + \frac{\rho_i}{\phi_i - \rho_i} \Pr\{\mathbf{W}_{-i}^{\rho_{-i}+\delta} > y\}.$$

Passing $y \to \infty$, and then $\delta \downarrow 0$, we obtain

$$\limsup_{x\to\infty} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{B}_i^r > x\}} \leq \frac{\rho_i}{1-\rho}. \tag{6}$$

Combining (5), (6) with Theorem 2.1, observing that $\frac{\rho_i}{1-\rho} = \frac{\rho_i}{c_i - \rho_i}$ completes the proof. □

4.2. On/Off processes

We finally characterize the tail distribution of the buffer content for sources with intermediately regularly-varying On-Off arrival processes. The classes of *regularly-varying* and *intermediately regularly-varying* distributions are denoted with the symbols $\mathcal{R}$ and $\mathcal{IR}$, respectively. The definitions of these classes are given in the appendix.

Theorem 4.2 *If $A_i^r(\cdot) \in \mathcal{IR}$ and $r_i > c_i$, then*

$$\Pr\{\mathbf{V}_i > x\} \sim \Pr\{\mathbf{W}_i^{c_i} > x\} \qquad \text{as} \quad x \to \infty,$$

with $c_i = 1 - \rho_{-i}$ representing the link rate minus the aggregate average rate of all other sources.

Theorem 4.2 says that the tail behavior of the buffer content of an individual source i (with intermediately regularly-varying On-periods) is equivalent to the tail behavior when that source is served in isolation at a *constant* rate which is equal to the link rate minus the aggregate average rate of all other sources. Thus, asymptotically, the buffer content of source i is only influenced by the On-period distribution of the other sources through their *aggregate average* rate. This is again somewhat similar to Theorem 10 in [11] and a more general result obtained by Agrawal et al. [2]. In particular, source i is largely insensitive to extreme activity patterns of other sources with 'heavier' On-period distributions. The precise values of the GPS weights are also inessential again, as long as they are larger than the traffic intensities (Assumption 2.1). The condition $A_i^r(\cdot) \in \mathcal{IR}$ in Theorem 4.2 can probably be somewhat relaxed. This is a topic for further study.

Proof of Theorem 4.2 The lower bound in Lemma 3.1 implies, for any $\delta > 0$ and y,

$$\frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} \geq \frac{\Pr\{\mathbf{W}_i^{1-\rho_{-i}+\delta} > x + y\}}{\Pr\{\mathbf{A}_i^r > \frac{x+y}{r_i-1+\rho_{-i}-\delta}\}} \frac{\Pr\{\mathbf{A}_i^r > \frac{x+y}{r_i-1+\rho_{-i}-\delta}\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} \Pr\{\mathbf{U}_{-i}^{\rho_{-i}-\delta} + \sum_{j \neq i} \mathbf{W}_j^{\phi_j} < y\}.$$

Using the fact that $\mathcal{IR} \subset \mathcal{S}$ and Theorem 2.2,

$$\liminf_{x\to\infty} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} \geq p_i \frac{\rho_i}{1-\rho+\delta} \Pr\{\mathbf{U}_{-i}^{\rho_{-i}-\delta} + \sum_{j\neq i} \mathbf{W}_j^{\phi_j} < y\} \liminf_{x\to\infty} \frac{\Pr\{\mathbf{A}_i^r > \frac{x}{r_i-1+\rho_{-i}-\delta}\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}}$$

Letting $y \to \infty$, then $\delta \downarrow 0$ and applying Definition 4, we arrive at

$$\liminf_{x\to\infty} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} \geq p_i \frac{\rho_i}{1-\rho}. \tag{7}$$

Similarly, the upper bound in Lemma 3.2 gives for any $\delta > 0$ and y, with $\delta < 1 - \rho$,

$$\begin{aligned} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} &\leq \frac{\Pr\{\mathbf{W}_i^{1-\rho_{-i}-\delta} > x - y\}}{\Pr\{\mathbf{A}_i^r > \frac{x-y}{r_i-1+\rho_{-i}+\delta}\}} \frac{\Pr\{\mathbf{A}_i^r > \frac{x-y}{r_i-1+\rho_{-i}+\delta}\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} \\ &+ \frac{\Pr\{\mathbf{W}_i^{\phi_i} > x\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} \Pr\{\mathbf{W}_{-i}^{\rho_{-i}+\delta} > y\}, \end{aligned}$$

so that

$$\begin{aligned} \limsup_{x\to\infty} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} &\leq p_i \frac{\rho_i}{1-\rho-\delta} \limsup_{x\to\infty} \frac{\Pr\{\mathbf{A}_i^r > \frac{x}{r_i-1+\rho_{-i}+\delta}\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} \\ &+ p_i \frac{\rho_i}{\phi_i - \rho_i} \Pr\{\mathbf{W}_{-i}^{\rho_{-i}+\delta} > y\}. \end{aligned}$$

Passing $y \to \infty$, then $\delta \downarrow 0$ and applying Definition 4, we obtain

$$\limsup_{x\to\infty} \frac{\Pr\{\mathbf{V}_i > x\}}{\Pr\{\mathbf{A}_i^r > x/(r_i - c_i)\}} \leq p_i \frac{\rho_i}{1-\rho}. \tag{8}$$

Matching (7), (8) with Theorem 2.2 then completes the proof.

□

To conclude the section, we briefly discuss the crucial role of Assumption 2.1. Intuitively, the assumption that $\rho_i < \phi_i$ for all $i = 1, \ldots, N$ has two major consequences: (i) source i is always guaranteed to receive service at a rate which is sufficient for stability, even when other sources generate large traffic bursts; (ii) when source i produces a large traffic burst, all other sources continue to be served at a stable rate, thus taking away a fraction $\sum_{j \neq i} \rho_j$ of the link rate. This provides an intuitive explanation of the result that an individual source with long-tailed traffic characteristics is effectively served at a constant rate which is equal to the link rate minus the aggregate average rate of all other sources.

Besides Assumption 2.1, we have only used the following two properties of the GPS discipline in the proofs of Theorems 4.1 and 4.2: (i) it is work-conserving; (ii) it guarantees minimum rates $\phi_1, \ldots, \phi_N$. Thus, the results do not rely on the specific way in which excess capacity is redistributed in GPS, but hold for any scheduling discipline with these two properties.

In fact, Assumption 2.1 is essential for Theorems 4.1 and 4.2 to hold. If the assumption that $\rho_i < \phi_i$ is relaxed, then two situations may arise: (i) when other sources generate large traffic bursts, source i may not receive service at a rate which is necessary for stability, and experience a significant queue build-up; (ii) when source i produces a large traffic burst, not all other sources may continue to be served at a stable rate, thus claiming *less* than a fraction $\sum_{j \neq i} \rho_j$ of the link rate.

In scenario (i), the tail behavior of the buffer content of source i may potentially be affected by sources with heavier tails, which drastically complicates the queueing analysis. In case (ii), exactly what rate the other sources will be able to take away from source i depends on the specific values of the GPS weights (or the detailed mechanics of the rate sharing scheme in general). This is the subject of a sequel paper.

Acknowledgment
The authors gratefully acknowledge a helpful observation made by Bert Zwart.

REFERENCES

1. Abate, J., Whitt, W. (1997). Asymptotics for M/G/1 low-priority waiting-time tail probabilities. *Queueing Systems* **25**, 173-233.
2. Agrawal, R., Makowski, A.M., Nain, P. (1998). On a reduced load equivalence for a fluid model under subexponential assumptions. Report INRIA Sophia-Antipolis.
3. Asmussen, S., Schmidli, H., Schmidt, V. (1997). Tail probabilities for non-standard risk and queueing processes with subexponential jumps. Preprint.
4. Bertsimas, D., Paschalidis, I.Ch., Tsitsiklis, J.N. (1997). Large deviations analysis of the generalized processor sharing policy. Preprint Boston University.
5. Bingham, N.H., Goldie, C.M., Teugels, J.L. (1987). *Regular Variation* (Cambridge University Press, Cambridge, UK).
6. Boxma, O.J., Cohen, J.W., Deng, Q. (1998). Heavy-traffic analysis of the M/G/1 queue with priority classes. In these proceedings.
7. Boxma, O.J. Dumas, V. (1997). Fluid queues with heavy-tailed activity period distributions. CWI Report PNA-R9705. In: *Computer Communications*, to appear.
8. Cohen, J.W. (1973). Some results on regular variation for distributions in queueing and fluctuation theory. *J. Appl. Prob.* **10**, 343-353.
9. Dupuis, P., Ramanan, K. (1998). A Skorokhod Problem formulation and large deviation analysis of a processor sharing model. *Queueing Systems* **28**, 109-124.

10. Jelenković, P.R. (1998). Asymptotic results for queues with subexponential arrival processes. In: *Self-similar Network Traffic and Performance Evaluation*, eds. K. Park, W. Willinger (Wiley, New York), to appear.
11. Jelenković, P.R., Lazar., A.A. (1996). Multiplexing on-off sources with subexponential on periods. *Adv. Appl. Prob.* **31**, to appear.
12. Jelenković, P.R., Lazar, A.A. (1998). Subexponential asymptotics of a Markov-modulated random walk with queueing applications. *J. Appl. Prob.* **35**, 325-347.
13. Jelenković, P.R., Lazar, A.A., Semret, N. (1997). The effect of multiple time scales and subexponentiality of MPEG video streams on queueing behavior. *IEEE J. Sel. Areas Commun.* **15**, 1052-1071.
14. Leland, W.E., Taqqu, M.S., Willinger, W. (1993). On the self-similar nature of Ethernet traffic. In: *Proc. SIGCOMM '93*, 183-193.
15. Massoulié, L. (1998). Large deviations for Polling and Weighted Fair Queueing service systems. Preprint France Télécom-CNET.
16. Pakes, A.G. (1975). On the tails of waiting-time distributions. *J. Appl. Prob.* **12**, 555-564.
17. Parekh, A.K., Gallager, R.G. (1993). A generalized processor sharing approach to flow control in integrated services networks: the single-node case. *IEEE/ACM Trans. Netw.* **1**, 344-357.
18. Zhang, Z.-L., Towsley, D., Kurose, J. (1995). Statistical analysis of the generalized processor sharing discipline. *IEEE J. Sel. Areas Commun.* **13**, 1071-1080.
19. Zwart, A.P. (1998). Sojourn times in a multiclass processor sharing queue. In these proceedings.
20. Zwart, A.P., Boxma, O.J. (1998). Sojourn time asymptotics in the M/G/1 processor sharing queue. CWI Report PNA-R9802.

Appendix Definitions

Let $F(\cdot)$ be a distribution function on $[0,\infty)$.

Definition 1 $F(\cdot)$ is called *long-tailed* ($F(\cdot) \in \mathcal{L}$) if

$$\lim_{x\to\infty} \frac{1-F(x-y)}{1-F(x)} = 1, \quad \text{for all real } y.$$

Definition 2 $F(\cdot)$ is called *subexponential* ($F(\cdot) \in \mathcal{S}$) if

$$\lim_{x\to\infty} \frac{1-F^{*2}(x)}{1-F(x)} = 2,$$

where $F^{*2}(\cdot)$ is the two-fold convolution of $F(\cdot)$ with itself.

Definition 3 $F(\cdot)$ is called *regularly varying of index* $-\nu$ ($F(\cdot) \in \mathcal{R}_{-\nu}$) if

$$F(x) = 1 - \frac{l(x)}{x^\nu}, \quad \nu \geq 0,$$

where $l : (0,\infty) \to (0,\infty)$ is a function of slow variation, i.e., $\lim_{x\to\infty} l(\eta x)/l(x) = 1$, $\eta > 1$.

Definition 4 $F(\cdot)$ is called *intermediately regularly varying* $(F(\cdot) \in \mathcal{IR})$ if

$$\lim_{\eta \downarrow 1} \liminf_{x \to \infty} \frac{1 - F(\eta x)}{1 - F(x)} \;=\; 1.$$

A key reference is [5]. It is well-known that $\mathcal{R} \subset \mathcal{IR} \subset \mathcal{S} \subset \mathcal{L}$. The class $\mathcal{R}$ contains the Pareto distribution, but not the Weibull, lognormal, Benktander distributions, which are in $\mathcal{S}$.