User-level performance of elastic traffic in a differentiated-services environment

S.C. Borst [a,b,c] R. Núñez-Queija [a,b] M.J.G. van Uitert [a]

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

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

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

Abstract

We consider a system with two service classes, one of which supports elastic traffic. The traffic characteristics of the other class can be completely general, allowing streaming applications as an important special case. The link capacity is shared between the two traffic classes in accordance with the Generalized Processor Sharing (GPS) discipline. GPS-based scheduling algorithms, such as Weighted Fair Queueing, provide a flexible mechanism for service differentiation and prioritization.

We examine the user-level performance of the elastic traffic. The elastic traffic users randomly initiate file transfers with a heavy-tailed distribution. Within the elastic traffic class, the active flows share the available bandwidth in an ordinary Processor-Sharing (PS) fashion. The PS discipline has emerged as a natural paradigm for evaluating the user-perceived performance of bandwidth sharing algorithms like TCP. For a certain parameter range, we establish that the transfer delay incurred by the elastic traffic flows is asymptotically equivalent to that in an isolated PS system with constant service rate. This service rate is only affected by the streaming traffic through its average rate. Specifically, the elastic traffic is largely immune from possible adverse traffic characteristics or performance degradation due to prioritization of the streaming traffic. This confirms that GPS-based multiplexing mechanisms achieve significantly better performance for both traffic classes than a static bandwidth partitioning approach.

Key words: delay asymptotics, differentiated services, Generalized Processor Sharing (GPS), heavy-tailed traffic, processor sharing

1 Introduction

The future Internet is expected to support a wide range of services on a common infrastructure, such as voice, video and data applications. The consolidation of several services on a single platform offers significant operational advantages. Besides the typical scaling efficiencies, a second benefit lies in the greater flexibility in supporting future applications whose features are intrinsically uncertain. While offering potential synergies, however, the co-existence of heterogeneous services also involves several challenging issues. In particular, different applications may not only have extremely diverse traffic characteristics, but also drastically different Quality-of-Service (QoS) requirements. The integration of heterogeneous services thus raises the need for differentiated QoS, catering to the specific requirements of the various traffic flows.

One potential approach to achieve service differentiation is through the use of discriminatory scheduling mechanisms, which distinguish between packets of the various traffic streams. Because of scalability issues, it is practically infeasible though to manipulate packets at the granularity level of individual flows in the core of any high-speed network. To avoid these complexity problems, traffic flows may instead be aggregated into a small number of representative classes, with scheduling mechanisms acting at the coarser level of composite streams. For example, the majority of applications may be broadly categorized into two service classes, one supporting *streaming* traffic, the other carrying *elastic traffic*. Streaming traffic is produced by audio and video applications for both real-time communication and reproduction of stored sequences (*'traces'*). Usually, the transmission rate has some intrinsic time profile, which may either be nearly constant or highly bursty, depending on the specific application. In both these cases, the QoS experienced by the users is mainly determined by integrity of the time profile, making small packet delay and low loss crucial requirements. Elastic traffic, on the other hand, results from the transfer of digital documents such as Web pages, files and e-mails. In contrast to streaming traffic, the transmission rate is continuously adapted over time, based on the level of congestion in the network. Typically, it is not so much the delay of individual packets that matters, but the total transfer delay of the document that determines the QoS perceived by the users.

The above notions are at the heart of the DiffServ proposal [3], which defines the EF class (Expedited Forwarding) for delay-sensitive traffic, and the AF class (Assured Forwarding) for traffic with some degree of delay tolerance. In view of the relative delay sensitivity, it is desirable that streaming applications receive some sort of priority over the elastic flows, at least over short time scales. The DiffServ philosophy postulates that coarse scheduling at the class level preempts the need for fine-grained scheduling at the level of individual flows, avoiding the scalability issues mentioned above. In particular, the queue for the streaming class should normally be so small that internal scheduling is not essential and simple FIFO queueing for example is adequate. The queue for the elastic class may be larger, but will be regulated by end-to-end flow control protocols.

Although some degree of priority for the streaming sessions is appropriate, strict priority scheduling may in fact not be ideal, since it may lead to starvation of the elastic traffic. Even temporary starvation effects may cause flow control protocols like TCP to suffer a severe degradation in throughput performance. The Generalized Processor Sharing (GPS) discipline provides a potential mechanism for implementing priority scheduling in a flexible way, with strict priority queueing as an extreme option [14,15]. In GPS-based scheduling algorithms, such as Weighted Fair Queueing, the link capacity is shared among the backlogged classes in proportion to certain class-defined weight factors. By setting the weight factor for the elastic class relatively low, one may provide some degree of priority to the streaming sessions, while avoiding starvation of the elastic traffic.

In the present paper we examine the user-level performance of the elastic traffic in the above situation, where the link capacity is shared between the two classes in a GPS manner. The elastic traffic users randomly generate files whose sizes follow a heavy-tailed distribution. The assumption of heavy-tailed file sizes is motivated by extensive measurement studies [7]. The other class could carry streaming applications as described above, but is in fact allowed to be completely general in terms of traffic characteristics, while the internal scheduling mechanism can be any work-conserving discipline. Within the elastic traffic class, the active users share the capacity in a standard processor-sharing (PS) fashion. The PS discipline has been commonly adopted as a convenient modeling abstraction for evaluating the user-perceived performance of bandwidth sharing algorithms like TCP [11,12].

The delay distribution in PS queues with heavy-tailed service requirements has attracted significant interest over the past few years. Under different distributional assumptions and by means of different proof methods, the tail of the delay distribution was shown to be asymptotically equivalent to that of the service requirement distribution [19,18,12,13,9]. (See Section 3 for a more detailed discussion.)

In parallel, the workload behavior in GPS queues with heavy-tailed traffic characteristics has been intensively studied, see for instance [5,6,10]. These papers showed that for a wide range of parameter settings the workload of an individual traffic class is asymptotically equivalent to that in an isolated system with a constant service rate. In many situations, the latter service rate equals the capacity of the original system reduced by the average rate of the other classes, hence the term *reduced-load equivalence* [1]. Subsequent work [17] generalized the reduced-load equivalence result to *networks* of GPS queues.

All of the above papers, however, focused on the *workload* behavior in GPS systems rather than delay asymptotics. To the best of our knowledge, the present paper is the first one to investigate the delay asymptotics for a PS queue within a GPS system. For a certain range of parameter values, we demonstrate that the transfer delay incurred by the elastic traffic flows is asymptotically equivalent to that in an isolated PS system with constant service rate. This service rate is only affected by the streaming traffic through its average rate. In particular, the elastic traffic is protected from possible adverse traffic characteristics or performance degradation due to prioritization of the streaming traffic. This confirms that GPS-based bandwidth sharing mechanisms produce better performance for both traffic classes than a static segregation scheme. A somewhat related paper is that of Tinnakornsrisuphap *et al.* [16].

The remainder of the paper is organized as follows. In Section 2, we present a detailed model description. We provide a discussion and interpretation of the main result in Section 3. The detailed proof may be found in Section 4. We conclude the paper with a brief summary in Section 5.

2 Model description

We first present a detailed model description. We consider two traffic classes sharing a link of unit rate. The link rate is shared in accordance with the Generalized Processor Sharing (GPS) discipline, which operates as follows. Class i is assigned a weight ϕ_i, $i = 1, 2$, with $\phi_1 + \phi_2 = 1$. As long as both classes are backlogged, class-i is served at rate ϕ_i. If one of the classes is not backlogged however, then the capacity is reallocated to the other class, which is then served at the full link rate (if backlogged). (In case of gradual input, it may occur that class 2 is not backlogged while generating traffic at some rate $r_2 < \phi_2$. In that case, only the *excess* capacity $\phi_2 - r_2$ is reallocated to class 2.)

Let $A_i(s,t)$ be the amount of class-i traffic generated during the time interval $(s,t]$. Denoting the class-i traffic intensity by ρ_i, we assume that

$$\lim_{t\to\infty} \frac{1}{t-s} A_i(s,t) = \lim_{u\to-\infty} \frac{1}{s-u} A_i(u,s) = \rho_i, \qquad \text{w.p. } 1, \tag{1}$$

for all s and $i = 1, 2$. The aggregate traffic process is denoted with $A(s,t) := A_1(s,t) + A_2(s,t)$ and the corresponding traffic intensity with $\rho := \rho_1 + \rho_2$. For stability, we assume that $\rho < 1$.

Class-1 customers arrive as a Poisson process of rate λ_1, and have service requirements with distribution $B_1(\cdot)$ and mean β_1, so that $\rho_1 = \lambda_1\beta_1$. (These additional assumptions are in agreement with (1).) We assume that the service requirements of class-1 customers are intermediately regularly varying, denoted as $B_1(\cdot) \in \mathcal{IR}$, which means that $\lim_{\alpha\uparrow 1} \limsup_{x\to\infty} \bar{B}_1(\alpha x)/\bar{B}_1(x) = 1$, with $\bar{B}_1(x) := 1 - B_1(x)$. Class-1 customers are served according to the Processor Sharing (PS) discipline, i.e., the available service rate for class 1 is equally shared among all the class-1 customers present. Thus, when there are n class-1 customers present at time t, each of them receives service at rate $c_1(t)/n$, with $c_1(t)$ representing the available service rate for class 1 at time t as governed by the GPS mechanism described above.

Other than (1), we do not make any specific assumptions regarding the traffic characteristics of class 2. Also, the internal scheduling mechanism for class 2 can be any work-conserving discipline.

We finally introduce some notation that will be used throughout the paper. For any two real functions $g(\cdot)$ and $h(\cdot)$, we write $g(t) = \mathrm{o}(h(t))$, as $t \to \infty$, if $\lim_{t\to\infty} g(t)/h(t) = 0$. We further use the notational convention $g(t) \sim h(t)$ to denote $\lim_{t\to\infty} g(t)/h(t) = 1$, or equivalently, $g(t) = h(t)(1 + \mathrm{o}(1))$ as $t \to \infty$.

3 Main result

We now formulate and explain the main result of the paper. To put things in perspective, we first review the sojourn time asymptotics in an ordinary (isolated) PS queue. Then we turn to the sojourn time asymptotics of class 1 in the integrated GPS system described in the previous section.

M/G/1 PS queue

Consider an ordinary isolated PS queue where class 1 is served at constant rate $c > \rho_1$. Let B_0 and S_0^c be the service requirement and sojourn time of a tagged customer arriving to the system at time 0. Denote by $W_1^c(t)$ the amount of class-1 work in the isolated system at time t, i.e., the sum of the remaining service requirements of all the customers in the system at time t except the tagged customer. Define $B_1^c(0, t)$ as the total amount of service received during the time interval $(0, t]$ by all customers except the tagged customer. Then B_0 and S_0^c satisfy the following identity relation

$$cS_0^c = B_0 + B_1^c(0, S_0^c),$$

with

$$B_1^c(0,t) = W_1^c(0) + A_1(0,t) - W_1^c(t), \qquad t \geq 0,$$

so that

$$cS_0^c = B_0 + W_1^c(0) + A_1(0, S_0^c) - W_1^c(S_0^c). \tag{2}$$

Large-deviations results suggest that rare events, given that they occur, almost exclusively tend to happen in the most probable manner. In the case of heavy-tailed processes, the most likely scenario for the rare event under consideration to occur often consists of a single extreme deviation in the traffic process. For a PS queue with a service requirement distribution of intermediate regular variation, the specific premise is that the most plausible cause for a large sojourn time S_0^c arises from a large service requirement B_0 of the customer itself, while the system otherwise shows average behavior. In particular, the amount of traffic generated over the course of the large sojourn time of the tagged customer is close to average, i.e., $A_1(0, S_0^c) \approx \rho_1 S_0^c$. In addition, the amount of work found upon arrival by the tagged customer is typical, and thus negligibly small compared to the service requirement B_0, i.e., $W_1^c(0) = \mathrm{o}(B_0)$. The amount of class-1 work that is left behind by the tagged customer is asymptotically negligible as well, i.e., $W_1^c(S_0^c) = \mathrm{o}(B_0)$. From (2), we thus obtain $cS_0^c \approx B_0 + \rho_1 S_0^c$, or equivalently, $S_0^c \approx B_0/(c - \rho_1)$. Of course, there are various other ways in which a large sojourn time may occur. In that sense, the above scenario only yields a lower bound for the probability of a large sojourn time occurring. However, all these alternatives are exceedingly improbable compared to the above scenario, so that the lower bound in fact provides the exact asymptotics, as reflected in the next theorem.

Theorem 1 *If $B_1(\cdot) \in \mathcal{IR}$ and $\rho_1 < c$, then*

$$\mathbb{P}\{\mathbf{S}_0^c > t\} \sim \mathbb{P}\{\mathbf{B}_0 > t(c - \rho_1)\}.$$

The above equivalence result was first proven for regularly varying service requirement distributions in Zwart & Boxma [19] using transform techniques. In a sequel paper of Zwart [18], the result was generalized to multi-class PS queues. Using a probabilistic proof based on conditional moments, Núñez-Queija [12,13] extended the tail equivalence result to PS models with a time-varying service capacity and intermediately regularly varying service requirements. Recently, Jelenković & Momčilović [9] used an alternative probabilistic method to generalize the result to a larger subclass of subexponential distributions with a so-called square-root insensitivity property. The latter class includes Weibull distributions with an index parameter smaller than 1/2. They further showed that the result is sharp, in the sense that the tail equivalence does not hold for Weibull distributions with a larger index parameter.

Remark 2 *The discussion preceding Theorem 1 in fact suggests that an extremely large service requirement and an extremely large sojourn time tend to occur simultaneously, which is a stronger statement than that of the theorem. This was indeed proved in [13, Lemmas 2.1 and 2.2] for an isolated PS queue. Although not explicitly stated, this stronger property also holds for the class-1 queue in the GPS system. This follows from the proofs in Section 4.*

Integrated GPS system

We now turn our attention to the class-1 sojourn time asymptotics in the integrated GPS system described in Section 2. Let B_0 and S_0 be the service requirement and sojourn time of a tagged class-1 customer arriving to the system at time 0. Denote by $W_i(t)$ the amount of class-i work in the system at time t, i.e., the sum of the remaining service requirements of all the class-i customers in the system at time t except the tagged class-1 customer. Define $B_i(0,t)$ as the total amount of service received during the time interval $(0,t]$ by all class-i customers except the tagged class-1 customer.

As before, the premise is that the most likely scenario for a large sojourn time S_0 to occur arises from a large service requirement B_0 of the tagged customer itself. However, the interaction with the competing traffic class must now be accounted for as well, as expressed in the following identity relation

$$S_0 = B_0 + B_1(0,S_0) + B_2(0,S_0), \tag{3}$$

with

$$B_i(0,t) = W_i(0) + A_i(0,t) - W_i(t), \qquad t \geq 0, \tag{4}$$

so that

$$S_0 = B_0 + W_1(0) + W_2(0) + A_1(0,S_0) + A_2(0,S_0) - W_1(S_0) - W_2(S_0). \tag{5}$$

As before, the amount of traffic generated by each class over the course of the large sojourn time S_0 is likely to be close to its average, i.e., $A_i(0,S_0) \approx \rho_i S_0$. In addition, the amount of work found upon arrival by the tagged customer is typical, and thus negligibly small compared to the service requirement B_0 since $\rho < 1$, i.e., $W_1(0)+W_2(0) = o(B_0)$. Further observe that $W_i(t) \leq W_i^{\phi_i}(t)$, since class i is guaranteed to receive a minimum service rate ϕ_i whenever backlogged (for class 2 this is an immediate consequence of the GPS mechanism; for class 1 the capacity taken away by the tagged customer needs to be taken into account, see Lemma 4 for a rigorous proof). Since $\rho_1 < \phi_1$, the amount of class-1 work that is left behind by the tagged customer is asymptotically negligible, i.e., $W_1(S_0) = \mathrm{o}(B_0)$. Similarly, the amount of class-2 work is negligible in case $\rho_2 < \phi_2$, i.e., $W_2(S_0) = \mathrm{o}(B_0)$. However, in case $\rho_2 > \phi_2$, class-2 work will accumulate over the course of the large sojourn time S_0 in a roughly linear manner at rate $\rho_2 - \phi_2$, i.e., $W_2(S_0) \approx (\rho_2 - \phi_2)S_0$, which may also be seen from $B_2(0, S_0) \approx \phi_2 S_0$. From (5), we thus obtain $S_0 \approx B_0 + (\rho_1 + \psi_2)S_0$, with $\psi_2 := \min\{\rho_2, \phi_2\}$, or equivalently, $S_0 \approx B_0/(1 - \psi_2 - \rho_1)$. Of course, there are various alternative scenarios in which a large sojourn time may occur, but these are all extremely unlikely compared to the above scenario.

The above heuristic arguments are confirmed by the next theorem, which is the main result of the paper. The detailed proof is provided in the next section, and proceeds by strengthening the above intuitive insight into rigorous statements.

Theorem 3 *If $B_1(\cdot) \in \mathcal{IR}$ and $\rho_1 < \phi_1$, then*

$$\mathbb{P}\{\mathbf{S}_0 > t\} \sim \mathbb{P}\{\mathbf{B}_0 > (1 - \psi_2 - \rho_1)t\}$$

and hence (by Theorem 1),

$$\mathbb{P}\{\mathbf{S}_0 > t\} \sim \mathbb{P}\{\mathbf{S}_0^{1-\psi_2} > t\}.$$

The above theorem shows that the class-1 sojourn time distribution is asymptotically equivalent to that in an isolated PS system with constant service rate $1 - \psi_2$. The latter service rate simply equals the original service rate reduced by either the average rate or the minimum guaranteed rate of class 2, whichever is lower. This phenomenon is reminiscent of the reduced-load equivalence result which was first established in Agrawal *et al.* [1] for the workload asymptotics of the superposition of a heavy-tailed On-Off source and a second lighter-tailed traffic process. In later work [6], this notion was generalized to the workload asymptotics in GPS queues with heavy-tailed traffic flows.

It is worth observing that the above theorem does not involve any specific assumptions regarding the traffic characteristics of class 2 other than through its average rate. In particular, the class-1 sojourn time asymptotics are not significantly affected by possible adverse traffic characteristics of class 2, provided $\rho_1 < \phi_1$. In case $\rho_1 > \phi_1$ (forcing $\rho_2 < \phi_2$ for stability), however, class 1 is no longer protected from class 2. In that case, class 1 will temporarily fail to receive a sufficiently large service rate so as to guarantee stability of class 1 when class 2 is backlogged, thus creating a potential alternative scenario for large sojourn times to occur. It strongly depends on the traffic characteristics of class 2 how likely this alternative scenario is compared to the earlier scenario. We conjecture that the above reduced-load equivalence result will in fact continue to hold as long as the traffic characteristics of class 2 are 'lighter' than those of class 1, namely when $\mathbb{P}\{\sup_{t\geq 0}\{A_2(0,t)-\phi_2 t\} > x\} = \mathrm{o}(\bar{B}_1(x))$ as $x \to \infty$, which is the most interesting case from a practical perspective. This is proved in [12, Chapter 5] for a situation where class 2 has strict priority over class 1 ($\phi_1 = 0$), assuming an infinite variance of the service requirements of class 1. If $\rho_1 > \phi_1$ and class 2 does not have lighter traffic characteristics (in the sense described above), the tail behavior of the class-1 sojourn time distribution is determined by the temporary starvation during activity periods of class 2, and thus inherits the traffic characteristics of class 2. This effect is similar to the concept of induced burstiness for the workload asymptotics in GPS queues as described in [6]. We expect that the derivation of the class-1 sojourn time asymptotics in this regime involves the analysis of unstable PS queues along the lines of Jean-Marie & Robert [8].

4 Proof of the main result

In this section we provide the proof of Theorem 3, by separately proving lower and upper bounds in Theorems 7 and 9, respectively. We first introduce the notion of a "permanent" class-1 customer, which plays a central role in the subsequent analysis of the sojourn time of an arbitrary tagged class-1 customer. Suppose that at time 0 a class-1 customer arrives, requiring an infinite amount of service. This customer will never leave the system and shares in the capacity as any ordinary class-1 customer. The amount of service received by the permanent customer over the interval $(0,t]$ is denoted with $B_0(0,t)$. Since the PS discipline does not discriminate between customers, irrespective of their service requirements, $B_0(0,t)$ would also be the amount of service received over the interval $(0,t]$ by the customer arriving at time 0 if it had had a finite service requirement, *provided that this service requirement is not less than* $B_0(0,t)$. Technically speaking, we may view the service process $B_0(0,t)$ of the (fictitious) permanent customer as a stochastic process and define the sojourn time $\mathbf{S}_0$ of a customer with service requirement $\mathbf{B}_0$ as a stopping time:

$$\mathbf{S}_0 := \inf\{t \geq 0 : \mathbf{B}_0 \leq B_0(0,t)\}.$$

Indeed, a customer's sojourn time exceeds t if and only if it requires an amount of service larger than that received during an interval of length t. We may therefore write

$$\mathbb{P}\{\mathbf{S}_0 > t\} = \mathbb{P}\{\mathbf{B}_0 > B_0(0,t)\}. \tag{6}$$

Paralleling (3) we may write the fundamental identity

$$B_0(0,t) + B_1(0,t) + B_2(0,t) = t, \tag{7}$$

where we use the fact that the permanent customer remains for ever in the system after time 0. The total backlog in the system at time t *not including the permanent customer* is denoted with $W(t) := W_1(t) + W_2(t)$. Because of (1) we also have, for all s,

$$\lim_{t\to\infty} \frac{1}{t-s} A(s,t) = \lim_{u\to-\infty} \frac{1}{s-u} A(u,s) = \rho, \qquad \text{w.p. } 1, \tag{8}$$

and since $\rho < 1$, the total backlog at time 0, expressed as

$$W(0) = \sup_{u\geq 0} \{A(-u,0) - u\},$$

has a non-defective distribution.

We construct an isolated reference system for class 1 by placing a copy of each arriving class-1 customer (with the same service requirement) in a PS queue with constant service capacity ϕ_1. At time 0, we also place a permanent customer in the reference queue. $W_1^{\phi_1}(t)$ and $B_0^{\phi_1}(0,t)$ represent the backlog due to non-permanent customers in the reference queue and the amount of service received by the permanent customer up to time $t \geq 0$, respectively. We assume that at time $-\infty$ both the original GPS and the reference system were empty.

Lemma 4 *At any time t each customer present in the GPS system has received at most the same amount of service in the reference system. As a consequence,*

$$W_1(t) \leq W_1^{\phi_1}(t), \tag{9}$$

and

$$B_0(0,t) \geq B_0^{\phi_1}(0,t) \tag{10}$$

for all t.

PROOF. First focus on a customer that arrives when there are no class-1 customers in either system. In the GPS system this customer will receive service at least at rate ϕ_1 until its departure or the next class-1 arrival, whichever occurs first. Clearly, during this period of time the assertions of the lemma hold.

Next, we show that this remains true for any time instant. Focus on a time $t = t_0$ for which the assertions of the lemma hold, i.e., all class-1 customers present in the GPS system have at least the same residual service requirement in the reference queue. (Besides these customers, there may be more customers present in the reference queue.) The total available service rate for class 1 in the GPS system (when backlogged) is never less than that in the reference system. At time t_0, the number of class-1 customers in the reference queue is not less than that in the GPS system. Therefore, the service rate for individual class-1 customers in the GPS system is not smaller than that in the reference system. Consequently, the assertions of the lemma remain true until the next arrival or departure (from either queue). Clearly, a class-1 arrival cannot invalidate the lemma, since an identical copy of the customer is placed in the reference queue. The same holds for a departure from either the GPS system alone or from both systems simultaneously, since a customer cannot obtain its total service requirement in the reference system sooner than in the GPS system. Note that the presence of the permanent customer after time 0 does not interfere with the above arguments. $\square$

The following is an immediate consequence of Lemma 4.

Corollary 5 *Let $\mathbf{S}_0$ and $\mathbf{S}_0^{\phi_1}$ represent the sojourn time of the same customer in the GPS system and in the reference system, respectively. Then, with probability 1,*

$$\mathbf{S}_0 \leq \mathbf{S}_0^{\phi_1}.$$

We now review several useful results from the literature. It is known [2, Section 8] that an isolated M/G/1 PS queue with a permanent customer is stable and, as $t \to \infty$, the distribution of $W_1^{\phi_1}(t)$ converges to that of a proper non-defective random variable $\mathbf{W}_1^{\phi_1}$:

$$\mathbb{P}\{\mathbf{W}_1^{\phi_1} \leq x\} := \lim_{t\to\infty} \mathbb{P}\{W_1^{\phi_1}(t) \leq x\}, \qquad x \geq 0.$$

For $c < \rho_2$, the following random variable represents the maximum of a random walk with a negative drift and, as such, has a proper non-defective distribution:

$$Z_2^c(s) := \sup_{u \geq s} \{c(u-s) - A_2(s,u)\}.$$

The next lemma gives a sample path upper bound for the service process $B_0(0,t)$, which in Theorem 7 results in a lower bound for the sojourn time.

Specifically, the lemma relates $B_0(0,t)$ to the average service rate available for the permanent customer, which equals $1-\rho_1-\psi_2$. The remaining terms account for random fluctuations around the mean.

Lemma 6 *For $t>0$ and $\epsilon>0$,*

$$B_0(0,t) \le (1-\rho_1-\psi_2+2\epsilon)t + (\rho_1-\epsilon)t - A_1(0,t) + W_1^{\phi_1}(t) + Z_2^{\psi_2-\epsilon}(0).$$

PROOF. Using (4), (7) and (9), we obtain

$$\begin{aligned} B_0(0,t) &\le t - A_1(0,t) + W_1(t) - B_2(0,t) \\ &\le t - (\rho_1-\epsilon)t + (\rho_1-\epsilon)t - A_1(0,t) + W_1^{\phi_1}(t) - B_2(0,t). \end{aligned}$$

Let s be the last time within the interval $[0,t]$ at which there was no backlog of class 2 in the GPS system. If such an s does not exist, take $s=0$. We obviously have

$$\begin{aligned} B_2(0,t) &\ge A_2(0,s) + \phi_2(t-s) \ \ge\ A_2(0,s) + \psi_2(t-s) \\ &\ge (\psi_2-\epsilon)t + A_2(0,s) - (\psi_2-\epsilon)s \ \ge\ (\psi_2-\epsilon)t - Z_2^{\psi_2-\epsilon}(0). \end{aligned}$$

This concludes the proof. □

Using Lemma 6, we prove the following probabilistic lower bound for the sojourn time.

Theorem 7 (lower bound) *If $B_1(\cdot) \in \mathcal{IR}$ and $\rho_1 < \phi_1$, then*

$$\liminf_{t\to\infty} \frac{\mathbb{P}\{\mathbf{S}_0 > t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2)t\}} \ge 1.$$

PROOF. Using (6) and Lemma 6, we have

$$\begin{aligned} &\mathbb{P}\{\mathbf{S}_0 > t\} \qquad (11)\\ &\ge \mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2+2\epsilon)t + (\rho_1-\epsilon)t - A_1(0,t) + W_1^{\phi_1}(t) + Z_2^{\psi_2-\epsilon}(0)\} \\ &\ge \mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2+4\epsilon)t\} \\ &\quad \times \mathbb{P}\{(\rho_1-\epsilon)t - A_1(0,t) + W_1^{\phi_1}(t) \le \epsilon t\}\, \mathbb{P}\{Z_2^{\psi_2-\epsilon}(0) \le \epsilon t\}. \end{aligned}$$

Note that $A_1(0,t)$ and $W_1^{\phi_1}(t)$ are not independent. Since $Z_2^{\psi_2-\epsilon}(0)$ has a non-defective distribution, it follows that

$$\lim_{t\to\infty} \mathbb{P}\{Z_2^{\psi_2-\epsilon}(0) \le \epsilon t\} = 1.$$

Observe that

$$
\begin{aligned}
&\mathbb{P}\{(\rho_1-\epsilon)t - A_1(0,t) + W_1^{\phi_1}(t) \le \epsilon t\}\\
\ge&\mathbb{P}\{A_1(0,t) \ge (\rho_1-\epsilon)t,\, W_1^{\phi_1}(t) \le \epsilon t\}\\
\ge&\mathbb{P}\{A_1(0,t) \ge (\rho_1-\epsilon)t\} - \mathbb{P}\{W_1^{\phi_1}(t) > \epsilon t\}.
\end{aligned}
$$

Now, $\mathbb{P}\{A_1(0,t) \ge (\rho_1-\epsilon)t\} \to 1$ as $t\to\infty$, because of (1). Moreover, for fixed $x>0$,

$$
\limsup_{t\to\infty} \mathbb{P}\{W_1^{\phi_1}(t) > \epsilon t\} \le \limsup_{t\to\infty} \mathbb{P}\{W_1^{\phi_1}(t) > x\} = \mathbb{P}\{\mathbf{W}_1^{\phi_1} > x\}.
$$

Letting $x\to\infty$ and using the fact that $\mathbf{W}_1^{\phi_1}$ has a proper distribution, we thus have

$$
\lim_{t\to\infty} \mathbb{P}\{W_1^{\phi_1}(t) > \epsilon t\} = 0.
$$

Using (11), we have

$$
\liminf_{t\to\infty} \frac{\mathbb{P}\{\mathbf{S}_0 > t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2+4\epsilon)t\}} \ge 1.
$$

Finally, use the fact that $B_1(\cdot) \in \mathcal{IR}$:

$$
\begin{aligned}
&\liminf_{t\to\infty} \frac{\mathbb{P}\{\mathbf{S}_0 > t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2)t\}}\\
\ge&\liminf_{t\to\infty} \frac{\mathbb{P}\{\mathbf{S}_0 > t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2+4\epsilon)t\}} \frac{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2+4\epsilon)t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2)t\}}\\
\ge&\liminf_{t\to\infty} \frac{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2+4\epsilon)t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2)t\}} \longrightarrow 1, \qquad \epsilon \downarrow 0. \qquad \square
\end{aligned}
$$

The following lemma implies an upper bound for the sojourn time. We will need this bound for the case $\rho_2 < \phi_2$. When $\rho_2 \ge \phi_2$, it turns out that the upper bound provided by Corollary 5 is already tight.

Lemma 8 *For $t \ge 0$,*

$$
B_0(0,t) \ge t - W(0) - A(0,t).
$$

PROOF. Directly from (4) and (7). $\square$

Corollary 5 and Lemma 8 provide the necessary ingredients to prove the next theorem which together with Theorem 7 implies the result of Theorem 3.

Theorem 9 (upper bound) *If $B_1(\cdot) \in \mathcal{IR}$ and $\rho_1 < \phi_1$ then*

$$\limsup_{t\to\infty} \frac{\mathbb{P}\{\mathbf{S}_0 > t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho_1-\psi_2)t\}} \leq 1.$$

PROOF. When $\rho_2 \geq \phi_2$, the theorem is immediately implied by Theorem 1 and Corollary 5. Let us therefore assume that $\rho_2 < \phi_2$, i.e., $\psi_2 = \rho_2$. By virtue of Corollary 5 and Lemma 8 we may write

$$\begin{aligned}
&\mathbb{P}\{\mathbf{S}_0 > t\} \\
\leq &\mathbb{P}\{\mathbf{S}_0^{\phi_1} > t,\, \mathbf{B}_0 > (1-\rho-\epsilon)t - W(0) + (\rho+\epsilon)t - A(0,t)\} \\
\leq &\mathbb{P}\{\mathbf{B}_0 > (1-\rho-2\epsilon)t\} \,+\, \mathbb{P}\{\mathbf{S}_0^{\phi_1} > t,\, W(0) + A(0,t) - (\rho+\epsilon)t > \epsilon t\} \\
\leq &\mathbb{P}\{\mathbf{B}_0 > (1-\rho-2\epsilon)t\} \,+\, \mathbb{P}\{\mathbf{S}_0^{\phi_1} > t,\, \mathbf{B}_0 \leq (\phi_1-\rho_1-\epsilon)t\} \\
&+ \mathbb{P}\{\mathbf{B}_0 > (\phi_1-\rho_1-\epsilon)t\}\, \mathbb{P}\{W(0) + A(0,t) - (\rho+\epsilon)t > \epsilon t\}.
\end{aligned}$$

Using [13, Lemma 2.1] — the assumptions of which are satisfied by the sojourn time in an isolated PS queue because of [13, Theorem 4.1] — we have

$$\mathbb{P}\{\mathbf{S}_0^{\phi_1} > t,\, \mathbf{B}_0 \leq (\phi_1-\rho_1-\epsilon)t\} = \mathrm{o}(\mathbb{P}\{\mathbf{B}_0 > (\phi_1-\rho_1-\epsilon)t\}), \quad t\to\infty.$$

Furthermore, as $t \to \infty$,

$$\begin{aligned}
&\mathbb{P}\{W(0) + A(0,t) - (\rho+\epsilon)t > \epsilon t\} \\
\leq &\mathbb{P}\{W(0) > \epsilon t\} + \mathbb{P}\{A(0,t) > (\rho+\epsilon)t\} \longrightarrow 0,
\end{aligned}$$

where we use (8) and the fact that $W(0)$ has a non-defective distribution. Therefore,

$$\begin{aligned}
\mathbb{P}\{\mathbf{S}_0 > t\} &\leq \mathbb{P}\{\mathbf{B}_0 > (1-\rho-2\epsilon)t\} \,+\, \mathrm{o}(\mathbb{P}\{\mathbf{B}_0 > (\phi_1-\rho_1-\epsilon)t\}) \\
&= \mathbb{P}\{\mathbf{B}_0 > (1-\rho-2\epsilon)t\} \,+\, \mathrm{o}(\mathbb{P}\{\mathbf{B}_0 > (1-\rho)t\}).
\end{aligned}$$

The last equality uses the fact that $B_1(\cdot) \in \mathcal{IR}$ and can be proved using [13, Appendix B]. Using $B_1(\cdot) \in \mathcal{IR}$ once more, we finally obtain

$$\limsup_{t\to\infty} \frac{\mathbb{P}\{\mathbf{S}_0 > t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho)t\}} \leq \limsup_{t\to\infty} \frac{\mathbb{P}\{\mathbf{B}_0 > (1-\rho-2\epsilon)t\}}{\mathbb{P}\{\mathbf{B}_0 > (1-\rho)t\}} \longrightarrow 1,$$

letting $\epsilon \downarrow 0$. □

5 Summary

We have investigated the delay of elastic traffic when served together with other traffic in a two-class GPS system. Within the elastic class, capacity is shared according to the PS mechanism, which ensures that bandwidth is equally divided among the customers present. Previous work has already shown that GPS-like mechanisms are able to protect flows in the sense that the *workload* asymptotics are not affected by other classes, unless the weights are chosen too small. Therefore, these mechanisms are considered relevant for future DiffServ applications.

We have shown that the delay asymptotics of the elastic traffic class are not affected by the other class, provided that the GPS weight of the elastic class is chosen larger than its average traffic rate.

The result established in Theorem 3 extends the known tail equivalence result for the ordinary M/G/1 processor-sharing queue (Theorem 1) to a GPS environment. In contrast, the sojourn time and service requirement distributions will not necessarily be equally heavy when the weight of the elastic class is too small, unless additional assumptions are imposed on the traffic characteristics of the other class (see e.g. [12] for a scenario with strict priority for the competing class).

References

[1] Agrawal, R., Makowski, A.M., Nain, Ph. (1999). On a reduced load equivalence for fluid queues under subexponentiality. *Queueing Systems* **33**, 5–41.

[2] Van den Berg, J.L., Boxma, O.J. (1991). The M/G/1 queue with processor sharing and its relation to a feedback queue. *Queueing Systems* **9**, 365–401.

[3] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., Weiss, W. (1999). A framework for differentiated services. IETF RFC 2475.

[4] Borst, S.C., Boxma, O.J., Jelenković, P.R. (1999). Induced burstiness in generalized processor sharing queues with long-tailed traffic flows. In: *Proc. 37th Annual Allerton Conference on Communication, Control, and Computing.*

[5] Borst, S.C., Boxma, O.J., Jelenković, P.R. (2000). Asymptotic behavior of generalized processor sharing with long-tailed traffic sources. In: *Proc. Infocom 2000 Conference*, Tel-Aviv, Israel, 912–921.

[6] Borst, S.C., Boxma, O.J., Jelenković, P.R. (2000). Reduced-load equivalence and induced burstiness in GPS queues with long-tailed traffic flows. CWI Report PNA-R0016. *Queueing Systems*, to appear.

[7] Crovella, M., Bestavros, A. (1996). Self-similarity in World Wide Web traffic: evidence and possible causes. In: *Proc. ACM Sigmetrics '96*, 160–169.

[8] Jean-Marie, A., Robert, Ph. (1994). On the transient behavior of the processor sharing queue. *Queueing Systems* **17**, 129–136.

[9] Jelenković, P.R., Momčilović, P. (2001). Resource sharing with subexponential distributions. Preprint, Columbia University.

[10] Kotopoulos, C., Likhanov, N., Mazumdar, R.R. (2001). Asymptotic analysis of the GPS system fed by heterogeneous long-tailed sources. In: *Proc. Infocom 2001 Conference*, Anchorage AK, USA, 299–308.

[11] Massoulié, L., Roberts, J.W. (1999). Bandwidth sharing: Objectives and algorithms. In: *Proc. Infocom 1999 Conference*, New York NY, USA, 1395–1403.

[12] Núñez-Queija, R. (2000). *Processor-Sharing Models for Integrated-Services Networks.* PhD Thesis, Eindhoven University of Technology, ISBN 90-646-4667-8.

[13] Núñez-Queija, R. (2002). Queues with equally heavy sojourn time and service requirement distributions. CWI Report PNA-R0201. *Ann. Oper. Res.*, to appear.

[14] 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.

[15] Parekh, A.K., Gallager, R.G. (1994). A generalized processor sharing approach to flow control in integrated services networks: the multiple node case. IEEE/ACM *Trans. Netw.* **2**, 137–150.

[16] Tinnakornsrisuphap, P., Agrawal, R., Feng, W.-C. (2001). Rate-adjustment algorithm for aggregate TCP congestion control. Preprint, University of Wisconsin.

[17] Van Uitert, M.J.G., Borst, S.C. (2001). Generalised Processor Sharing networks fed by heavy-tailed input flows. In: *Proc. Infocom 2001 Conference*, Anchorage AK, USA, 269–278.

[18] Zwart, A.P. (1999). Sojourn times in a multiclass processor sharing queue. In: *Teletraffic Engineering in a Competitive World, Proc. ITC-16*, Edinburgh, UK, eds. P. Key, D. Smith (North-Holland, Amsterdam), 335–344.

[19] Zwart, A.P., Boxma, O.J. (2000). Sojourn time asymptotics in the M/G/1 processor sharing queue. *Queueing Systems* **35**, 141–166.