Delay asymptotics in the $GI/GI/1$ PS queue

Bert Zwart

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

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

zwart@win.tue.nl

In this talk, we consider the $GI/GI/1$ PS (Processor Sharing) queue with load $\rho < 1$ in steady state. In particular, we investigate the tail behavior of $P(V > x)$ as $x \to \infty$ with V being the steady-state sojourn time of a customer.
We first concentrate on the case where the service time B has a heavy-tailed distribution. For the $M/G/1$ PS queue, several studies (see e.g. [5, 6]) have shown that, under certain assumptions,

$$P(V > x) \sim P(B > x(1 - \rho)). \tag{1}$$

All proofs of (1) depend on the fact that the queue lenght distribution in steady state is geometric. Since it is even unknown whether the steady-state $GI/GI/1$ PS queue lenght can be bounded by a geometric tail, we develop a different method based on the cycle formula for regenerative processes. Under some additional assumptions, this results in extensions of (1) to the $GI/GI/1$ queue and even the multiclass Discriminatory Processor Sharing queue.
Next, we move on to the case where service-time distributions are light-tailed. This is an interesting problem in itself, given the following remarkable asymptotic for the $M/M/1$ PS queue [4, 1]:

$$P(V > x) \sim ax^{-5/6}\mathrm{e}^{-bx^{1/3}}\mathrm{e}^{-\gamma x}.$$

We supplement this result by proving logarithmic asymptotics for the $GI/GI/1$ queue, i.e., we prove a result of the form

$$P(V > x) \sim \gamma x. \tag{2}$$

The proof of this result combines a change of measure argument with a recent fluid limit result for overloaded PS queues, cf. [8].
Time permitting, we conclude by giving some partial results for a PS queue where two classes of customers (with either light-tailed or heavy-tailed (Pareto) service times) arrive. An interesting question is whether the tail of the sojourn time of the light tailed customers $P(V_L > x)$ depends on the heavy-tailed class. We give a lower bound for $P(V_L > x)$ which indicates that $P(V_L > x)$ is not light-tailed or Pareto, but Weibullian. This implies that the effect of heavy-tailed customers is significant, but not as significant as in FCFS queues.

The above results can be found in the preprints [3, 7, 2], which are available upon request.

References

[1] Borst, S.C., Boxma, O.J., Morrison, J.A., Núñez-Queija, R. (2003). The equivalence between processor sharing and service in random order. *Operations Research Letters* **31**, 254–262.

[2] Borst, S.C., Núñez-Queija, R., Zwart, A.P. Bandwidth sharing with heterogeneous flow sizes. To appear in *Annals of Telecommunications*.

[3] Borst, S.C., van Oothegem, D. Zwart, A.P. (2004). Tail asymptotics for discriminatory processor sharing queues with heavy-tailed service sequirements. To appear in *Performance Evaluation* (special edition on heavy tails and LRD).

[4] Flatto, L. (1997). The waiting time distribution for the random order service $M/M/1$ queue. *Annals of Applied Probability* **7**, 382–409.

[5] Guillemin. F., Robert, P., Zwart, A.P. (2004) Asymptotic results for Processor Sharing queues. *Advances in Applied Probability*, june 2004.

[6] Jelenković, P., Momčilović, P. (2003). Large deviation analysis of subexponential waiting times in a processor-sharing queue. *Mathematics of Operations Research* **28**, 587–608.

[7] Mandjes, M., Zwart, A.P. (2004). Large deviations and importance sampling for sojourn times in Processor Sharing queues. Submitted to *Queueing Systems*.

[8] Puha, A., Stolyar, A., Williams, R. (2004). The fluid limit of an overloaded processor sharing queue. Preprint.