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

Bert Zwart
Department of Mathematics & Computer Science
Eindhoven University of Technology
HG 9.15, P.O. Box 513
5600 MB Eindhoven, The Netherlands

Abstract

In this talk, we consider $GI/GI/1$ 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 $GI/GI/1$ 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 distribution in steady state is geometric. Since it is even unknown whether the steady-state $GI/GI/1$ PS queue length 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]:

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

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] S. C. Borst, O. J. Boxma, J. A. Morrison, and R. Núñez-Queija, *The equivalence between processor sharing and service in random order*, Oper. Res. Lett. **31** (2003), no. 4, 254–262.

[2] S. C. Borst, R. Núñez-Queija, and A. P. Zwart, *Bandwidth sharing with heterogeneous flow sizes*, Ann. Telecomm. (to appear).

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

[4] L. Flatto, *The waiting time distribution for the random order service $M/M/1$ queue*, Ann. Appl. Probab. **7** (1997), no. 2, 382–409.

[5] F. Guillemin, P. Robert, and A. P. Zwart, *Asymptotic results for Processor Sharing queues*, Adv. in Appl. Probab. (2004).

[6] P. Jelenković and P. Momčilović, *Large deviation analysis of subexponential waiting times in a processor-sharing queue*, Math. Oper. Res. **28** (2004), 587–608.

[7] M. Mandjes and A. P. Swart, *Large deviations and importance sampling for sojourn times in Processor Sharing queues*, Queueing Syst. (submitted).

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