On the Variance of the Least Attained Service Policy and its Use in Multiple Bottleneck Networks

Matthias Auchmann[1] and Guillaume Urvoy-Keller[2]

[1] Technische Universität Wien, Austria
m.auchmann@artech.at
[2] Eurecom,France
Guillaume.Urvoy@eurecom.fr

Abstract. Size-based scheduling has proved to be effective in a lot of scenarios involving Internet traffic. In this work, we focus on the Least Attained Service Policy, a popular size-based scheduling policy. We tackle two issues that have not received much attention so far. Firstly, the variance of the conditional response time. We prove that the classification proposed by Wierman et al. [11], which classifies LAS as an always unpredictable policy, is overly pessimistic. We illustrate the latter by focusing on the M/M/1/LAS queue. Secondly, we consider LAS queues in tandem. We provide preliminary results concerning the characterization of the output process of an M/M/1/LAS queue and the conditional average response time of LAS queues in tandem.

1 Motivation

Size-based scheduling has proved to be very effective in increasing performance in a lot of scenarios: Web servers [4], Internet traffic [8] or 3G networks [6]. The key idea behind size-based scheduling is to favor short jobs while ensuring that large jobs do not starve. The net result is better interactivity from the user point of view as short jobs correspond to interactive applications, while large job correspond to bulk transfers when considering Internet traffic. The extent to which large jobs suffer depends on the statistical characteristics of the job size distribution and especially onhow the mass is distributed among short and large jobs. Broadly speaking, the larger the mass carried by the large flows, the smaller the penalty since short flows, that have the highest priority, can not monopolize the server. Heavy-tailed distributions, which have often been observed in the Internet [3], feature such a property.

In this paper, we consider the Least Attained Service (LAS) policy, a.k.a the Foreground-Background policy [7]. LAS has been initially proposed and studied in the context of time-sharing computers in the late 60s [10]. Under LAS, priority is given to the job that has received the least amount of service. In case of ties, jobs share the server in a round-robin manner. A salient feature of LAS is that it has no internal parameter to tune.

Our focus in this work is twofold. First, we focus on the conditional variance of LAS. Few results exist concerning the variance of LAS [11, 2]. In [2], the asymptotic conditional variance is considered. In [11], the authors propose a classification of scheduling policies based on their variance. In particular, they propose to classify scheduling policies as: (i) always predictable, (ii) sometimes predictable or (iii) always unpredictable - precise definitions will be given in Section 2. LAS falls in the latter category, which seems at first sight disappointing. Indeed, LAS does a good job at providing low response time to small jobs but despite this nice property, the results in [11] seem to restrict the interest of LAS. We revisit the variance of LAS by considering the case of an M/M/1/LAS queue. This case is interesting for two reasons. First, it is mathematically tractable. Second, the performance of LAS and especially the fraction of flows that receive a better service under LAS than under PS is known to increase with the variability of job size distribution [8] ; and empirical distributions observed for Internet traffic have a much higher variability than the one of an exponential distribution [5].
Considering the M/M/1/LAS queue, we analytically bound the fraction of flows that are treated in a predictable manner. We obtain that at least 75% of flows are treated predicably, irrespectively of the load. Numerical studies further demonstrate that the actual fraction of such flows should be closer to 95%.

Second, we focus on the problem of using LAS queues in tandem. The motivation behind this scenario is to determine the benefits that could be obtained with LAS in a multiple bottlenecks scenario. A typical example is wireless mesh networks based on the 802.11 protocol where the available bandwidth is known to be highly varying, as has been exemplified by the roofnet experiment (http://pdos.csail.mit.edu/roofnet/doku.php). In such a situation, LAS could be highly beneficial as it allows to maintain a minimum level of interactivity, even when congestion is high. However, using LAS at multiple queues in tandem can also be detrimental to large flows that could be penalized multiple times.

To the best of our knowledge, no work has tackled the problem of studying LAS in tandem queues. We rely on numerical evaluations to address this issue as the analytical approach seems too complex at the moment, even for the case of M/M/1/LAS queues. We make the following contributions: (i) We demonstrate that while the departure process of an M/M/1/LAS queue is apparently a Poisson process, the LAS scheduling policy introduces a negative correlation between departures times and job sizes; (ii) We evaluate the impact of the above correlation on the conditional average response times of LAS queues in tandem.

2 Conditional Variance of an M/M/1/LAS queue

The variance of the conditional response time for an $M/G/1/LAS$ queue with arrival rate λ is given by [12]:

$$Var[T(x)]^{LAS} = \frac{\lambda x \tilde{m}_2(x)}{(1-\tilde{\rho}(x))^3} + \frac{\lambda \tilde{m}_3(x)}{3(1-\tilde{\rho}(x))^3} + \frac{3}{4}\Big(\frac{\lambda \tilde{m}_2(x)}{(1-\tilde{\rho}(x))^2}\Big)^2 \qquad (2.1)$$

Where $\{\tilde{m}_i(x)\}_{i\geq 1}$ are the truncated moments of the service time distribution. Truncated moments converge to the moments $\{m_i\}_{i\geq 1}$ when the job size tends to infinity. In particular, $\tilde{m}_1(\infty) = m_1 = \frac{1}{\mu}$ is the mean job service time. $\tilde{\rho}(x) = \lambda\tilde{m}_1(x)$ is the (truncated) load of the jobs up to size x, where $\tilde{\rho}(\infty) = \rho = \frac{\lambda}{\mu}$ is the load of the $M/G/1/LAS$ queue. Truncated moments for an exponential distribution $Exp(\mu)$ are given by:

$$\tilde{m}_i(x) = \int_0^x y^i \mu e^{-\mu y} dy + x^i e^{-\mu x}$$

In [11], scheduling policies were classified based on the variance of conditional response times as always predictable, always unpredictable or sometimes predictable. For a policy P, jobs of size x are treated predictably if:

$$\frac{Var[T(x)]^P}{x} \leq \frac{\lambda m_2}{(1-\rho)^3} \tag{2.2}$$

Otherwise jobs of size x are treated unpredictably. See [11] for a justification of the right side term in Eq. (2.2). A policy is predictable for a given load ρ and a given service time distribution if all job sizes are treated predictably. More generally, a scheduling policy P is:

- **Always predictable** if it is predictable under all loads and service distributions;
- **Sometimes predictable** if it is predictable under some loads and service distributions, and unpredictable under others;
- **Always unpredictable** if it is unpredictable regardless of service distribution and load.

In [11] a variance bound was derived to show that LAS is always unpredictable. In contrast, PS is shown to be always predictable. Comparison of LAS to PS is important as it is shown in [9] that the M/G/1/LAS queue is an accurate model for a LAS router while the M/G/1/PS queue is an accurate model for a FIFO router for connections with homogeneous RTT.

Our initial motivation in this paper is to show that LAS outperforms PS not only in terms of conditional response time offered to a majority of short jobs [8], but also in terms of conditional variance. Figure 1 illustrates the relative performance of LAS and PS for an M/M/1 queue with $\lambda = 1$ and $\mu = 1.25$, i.e. a load $\rho = 0.8$. We observe that while LAS offers both low average and variance of response time for most of the jobs, its performance becomes eventually worse than PS for the largest jobs.

The large variance observed for the large jobs in Figure 1 illustrates why LAS is classified as always unpredictable, but it is for the largest jobs only that Eq. (2.2) is violated. The question we address here is to determine the fraction of jobs that are treated predictably under LAS for the case of an M/M/1/LAS queue.

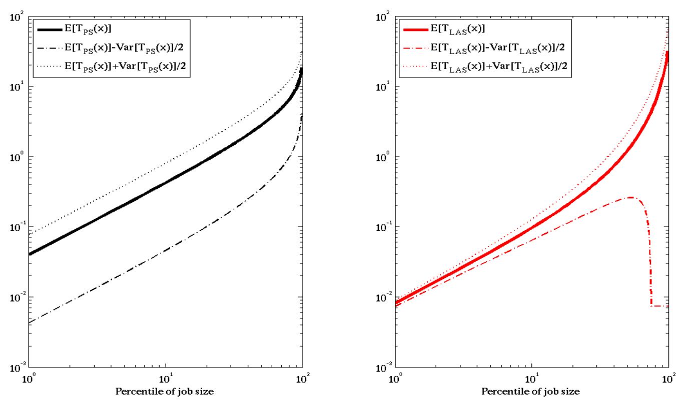

Fig. 1. M/M/1/LAS against M/M/1/PS: mean and variance of response time

2.1 M/M/1/LAS variance bounds

In this section, we first present two bounds on the variance of the conditional response time for an M/M/1/LAS queue. We next derive corresponding bounds on the fraction of jobs that are treated predictably under LAS. Due to space constraint, proofs are omitted, but can be found in our technical report [1].

Bound 1: For M/M/1/LAS,

$$Var[T(x)]^{LAS} \leq \frac{\lambda x m_2}{(1-\tilde{\rho}(x))^4}\Big(1+\frac{2}{e^2}-(\frac{1}{4}+\frac{2}{e^2})\tilde{\rho}(x)\Big) \tag{2.3}$$

Bound 2: For M/M/1/LAS,

$$Var[T(x)]^{LAS} \leq \frac{8e\rho+9\rho^2-8e\rho^2}{3e^2(1-\rho)^4}x^2 \tag{2.4}$$

Bounds 1 and 2 complement each other in the sense that bound 1 is more accurate for large job sizes while bound 2 is more accurate for small job sizes. Using these two bounds we further derive the percentage of jobs which are treated predictably in an M/M/1/LAS queue.

Bound 3: For M/M/1/LAS and a given ρ, at least a fraction of $\frac{1-\left((1-\rho)^3(1+\frac{2}{e^2})\right)^{1/4}}{\rho}$ of the jobs is treated predictably.

Bound 4: For M/M/1/LAS and a given ρ, at least a fraction of $1-e^{-\frac{6e^2(1-\rho)}{8e+9\rho-8e\rho}}$ of the jobs is treated predictably.

Note that bounds 3 and 4 depend on ρ only, not on λ and μ. Bounds 3 and 4 show that the majority of jobs are in fact treated predictably in a $M/M/1/LAS$

queue, as can be seen in Figure 2 where we plot the maximum of the two bounds (solid line). The minimum job percentage that is guaranteed to be treated predictably in this case, is about 75 percent.

As the above approach relies on lower bounding the number of jobs that are treated unpredictably under LAS, we further evaluated the percentage of jobs treated predictably using simulations. Figure 2 reports the results obtained by a simulator written in Matlab. Each simulation involved more than 500,000 jobs, and 4 different factorizations of ρ by using four different (λ, μ) pairs, which is why there are 4 points for each ρ value. We observe from Figure 2 that the fraction of flows that are treated predictably seems to depend on ρ only, not on the exact factorization of ρ. In addition, one can see that the actual percentage of jobs being treated predictably is very high: at least 90 percent of the jobs are treated predictably, regardless of the actual load value.

Note that for the case of Internet traffic, observed job size distributions are in general more skewed than the exponential one. As a consequence, we can expect that the fraction of jobs treated predictably is even higher than for the exponential case. As an illustration, for a Lognormal distribution with a coefficient of variation of 2, we observed that at all the load values we tested, between 10 and 90%, at least 99% of the jobs were treated predictably.

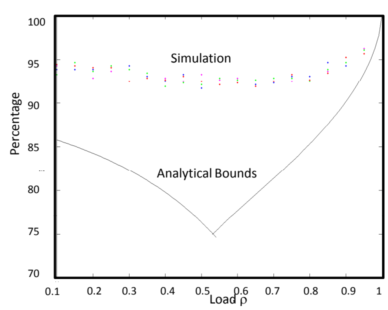

Fig. 2. Bounds 3 and 4 (solid line) and simulations (dots)

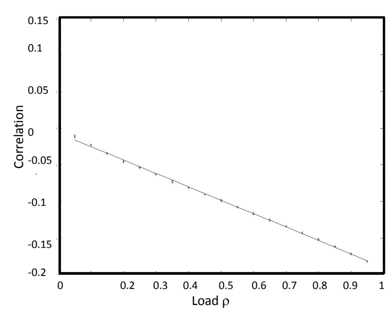

Fig. 3. Correlation of interdeparture times and jobsize for LAS

3 LAS in tandem queues

In this section, we focus on LAS in tandem queues where fresh arrivals occur only at the first queue and follow a Poisson process and service requirements follow an exponential distribution. We first study the output process of an isolated LAS queue and then, we will discuss the impact of the characteristics of the output process to the total response time of two LAS queues in tandem. Note that, to the best of our knowledge, no work has tackled this issue so far.

3.1 Characterization of the output process of an M/M/1/LAS queue

LAS being a work-conserving and blind scheduling policy (job sizes are not known in advance), Burke's theorem is valid for an M/M/1/LAS queue. Hence, the output process for this queue is a Poisson process. We however show in the next section that job sizes and inter-departure times become correlated at the output of an M/M/1/LAS queue.

Correlation of Inter-departure Time and Job size To further characterize the output process of an M/M/1/LAS queue, we investigated whether job size is independent of interdeparture time. Figure 3 shows the correlation between job sizes and interdeparture times for LAS, plotted for different utilizations ρ. Again, different (λ, μ) pairs are used for the plots, but since curves overlap, correlation seems to depend on ρ only. Correlation is of negative type and increases with increasing load. We believe that this correlation stems from the fact that as load increases, large jobs are interrupted more frequently by shorter jobs. As a consequence, their remaining service requirement can reach very low values, lower than the service requirements of short jobs. When a large job leaves the queue, it is highly likely to be in a short period of time where load is low. During this period, many large jobs that were stuck in the queue, with small remaining service times, leave the system. We believe this explains the negative correlation we observe.

Note that we did not observe such a correlation for the PS queuing discipline (we do not present results here due to space constraint).

3.2 Impact on tandem queue performance

The negative correlation observed above relates to the behavior of LAS that tends to sort jobs in the queue in ascending order, which means that short jobs tend to leave the LAS queue in groups and the same for large jobs. When considering the case of a tandem queues, this sorting introduced by the scheduler can have a detrimental impact on the performance of large flows. Indeed, when large flows reach the second queue, they again have to compete with the large flows they were in competition with in the first queue.

We quantitatively evaluated the previous intuition by comparing the conditional average response times[3] of two LAS queues in tandem with the sum of response time for the same system where we "re-draw" the job sizes of the job leaving the first queue using the initial distribution, thus wiping out the correlation observed in the previous section.

Figure 4 illustrates the above scenario for the case where load is 0.8. The left plot of Figure 4 represents the actual measured response time of the tandem system versus the added theoretical values (called independent queues in the graph). The right plot depicts the relative difference of the two systems.

[3] Results on the variance of the response time can be found in [1].

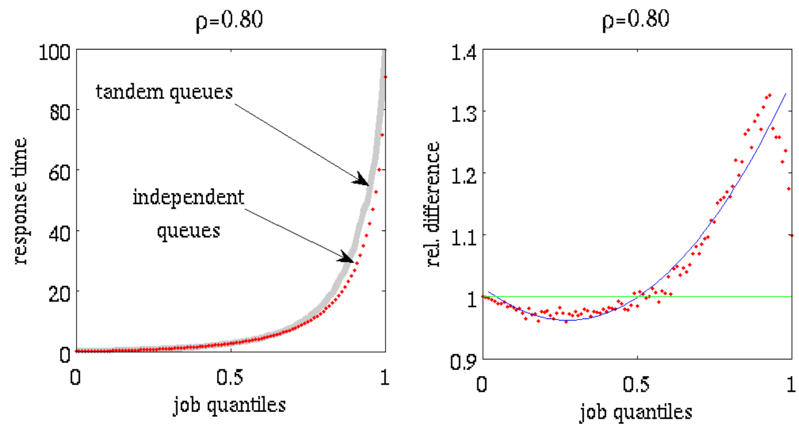

Fig. 4. Tandem Queue simulation results for LAS, $\rho = 0.80$

One can observe from Figure 4 that the observed negative correlation has a strong effect on the plots. While the response time is better than in the independent model for small jobs, LAS further penalizes large jobs in the tandem system.

4 Discussion

In this paper, we have focused both on the variance of the conditional response time of an M/M/1/LAS queue and the use of LAS in tandem queues.

For the variance, we proved that the fraction of jobs that are treated predictably is very high even in the unfavorable case of exponential service requirements. The practical implication of this result is that it is likely that when LAS is used for real Internet traffic, the variance in response time it offers to the majority of the flows be small as compared to the legacy FIFO scheduling policy.

Concerning the use of LAS in tandem queues, we demonstrated that while the output process of an M/M/1/LAS queue is Poisson, LAS tends to group jobs of similar sizes together, which results in a high penalty for the large jobs that cross the two queues. Practical implications of this result are less clear than for the variance case. Indeed routers work on a packet basis and do not output all the packets of a connection simultaneously as it happens in an M/M/1/LAS queue. This should dampen the effect we have observed.

As future work, we plan to continue working on the use of LAS in wireless mesh networks to precisely assess its performance. Note that a typical path in a wireless mesh network from an end user to an Internet gateway should be quite small, no more than 4 hops. As a consequence, we can expect that the nice properties of LAS, namely its ability to maintain interactivity even when the load is high, outweighs its side effect, namely the penalty experienced by large flows crossing multiple bottlenecks.

References

1. M. Auchmann and G. Urvoy-Keller. On the variance of the least attained service policy and its use in multiple bottleneck networks. Technical Report RR-08-224, Institut Eurecom, France, June 2008, http://www.eurecom.fr/util/popuppubli.fr.htm?page=copyright&id=2497.
2. P. Brown. Comparing fb and ps scheduling policies. Technical report, Orange Labs, 2008.
3. M. E. Crovella et al. *A Practical Guide to Heavy Tails*, chapter 3. Chapman and Hall, New-York, 1998.
4. M. Harchol-Balter, B. Schroeder, N. Bansal, and M. Agrawal. Size-based scheduling to improve web performance. *ACM Trans. Comput. Syst.*, 21(2):207–233, 2003.
5. F. Hernandez-Campos, M. Karaliopoulos, M. Papadopouli, and H. Shen. Spatio-temporal modeling of traffic workload in a campus wlan. In *WICON '06*, page 1, 2006.
6. M. Hu, J. Zhang, and J. Sadowsky. Size-aided opportunistic scheduling in wireless networks. In *GLOBECOM '03. IEEE*, 2003.
7. M. Nuyens and A. Wierman. The foreground-background queue: A survey. *Perform. Eval.*, 65(3-4):286–307, 2008.
8. I. A. Rai, G. Urvoy-Keller, and E. W. Biersack. Analysis of las scheduling for job size distributions with high variance. In *Proc. ACM SIGMETRICS*, pages 218–228, June 2003.
9. I. A. Rai, G. Urvoy-Keller, M. Vernon, and E. W. Biersack. Performance models for las-based scheduling disciplines in a packet switched network. In *ACM SIGMETRICS-Performance*, June 2004.
10. L. E. Schrage. The queue m/g/1 with feedback to lower priority queues. *Management Science*, 13(7):466–474, 1967.
11. A. Wierman and M. Harchol-Balter. Classifying scheduling policies with respect to higher moments of conditional response time. *ACM SIGMETRICS Performance Evaluation Review*, 33(1):229–240, June 2005.
12. S. Yashkov. Processor-sharing queues: some progress in analysis. *Queueing Systems: Theory and Applications*, 2(1):1–17, June 1987.