The Average Response Time in a Heavy-traffic SRPT Queue

Minghong Lin
CMS, California Institute of Technology

Adam Wierman
CMS, California Institute of Technology

Bert Zwart
CWI Amsterdam, VU University Amsterdam, Eurandom & Georgia Tech.

ABSTRACT

Shortest Remaining Processing Time first (SRPT) has long been known to optimize the queue length distribution and the mean response time (a.k.a. flow time, sojourn time). As such, it has been the focus of a wide body of analysis. However, results about the heavy-traffic behavior of SRPT have only recently started to emerge. In this work, we characterize the growth rate of the mean response time under SRPT in the M/GI/1 system under general job size distributions. Our results illustrate the relationship between the job size tail and the heavy traffic growth rate of mean response time. Further, we show that the heavy traffic growth rate can be used to provide an accurate approximation for mean response time outside of heavy traffic.

1. INTRODUCTION

Shortest Remaining Processing Time first (SRPT) has long been known to optimize the mean response time (a.k.a. flow time, sojourn time) in a single server queue [18]. As a result, there has been extensive research studying SRPT in a wide variety of models over the last 50 years [19, 16, 7]. Further, there has been renewed interest in SRPT recently as a result of a number of computer system designs based on SRPT-like policies [9, 17, 11]. This renewed interest has led to new results studying the tail behavior [4], the heavy-traffic behavior [3, 2], and the fairness of SRPT [21].

However, despite this large literature, there are some simple properties of SRPT that are still not well understood. One such property is the focus of this paper: *How does the mean response time under SRPT scale in heavy-traffic?*

It is perhaps surprising that this question is still not understood given its fundamental nature, especially since the mean response time, $E[T]$, was derived for the first time by Schrage & Miller [19] in 1966. However, the formula for the mean response time is complicated enough that the dependence of it on load, ρ, is not well understood. Specifically, there are a few papers that have derived the heavy-traffic growth rate of SRPT under specific job size distributions [2, 3, 15]. However, there is no complete characterization of the heavy-traffic growth rate of $E[T]$ under general job size distributions.

The contribution of the current paper is to provide a characterization of the heavy-traffic growth rate of $E[T]$ in an M/GI/1 preempt-resume queue (see Theorems 2, 3 and 4). This characterization highlights the relationship between the growth rate and the tail of the job size distribution. Specifically, the heavy-traffic growth rate is shown to depend on the tail of a measure $G(x)$, which characterizes the truncated load, and the Matuszewska index of the job size distribution, which relates to the moment conditions of the job sizes. The results illustrate that SRPT provides an order of magnitude of improvement over other common scheduling policies such as Processor Sharing (PS) and First Come First Served (FCFS) if and only if the job size distribution is unbounded. Further, the results illustrate that a heavier-tail implies a slower growth rate. Additionally, once the tail is "heavy-enough" (i.e. job sizes have an infinite variance) the growth rate becomes (up to a constant) independent of the job size distribution.

In addition to the insight provided by the heavy-traffic growth rate of SRPT, we illustrate that the heavy-traffic analysis can be used to provide a simple approximation of $E[T]$ under SRPT, which is accurate even outside of heavy-traffic. This simple approximation is useful when analyzing more complex models which have pieces that use SRPT. For example, this approximation has already been applied to attain results for a multi-queue load balancing model [5] and a speed scaling model [1]. Finally, the characterization of the growth rate of SRPT is especially important because of the optimality of SRPT. The results in this paper provide a baseline with which to compare the performance of other policies.

2. PRELIMINARIES

We study the performance of SRPT in an M/GI/1 preempt resume queue. Under SRPT at every instant, the job with the smallest remaining service time is scheduled. We assume the c.d.f. of job sizes, $F(x)$, is continuous. Denote by $E[T]$ the mean response time (a.k.a. sojourn time) under SRPT, which is the time from when a job enters the system until it completes service. Let $\bar{F}(x) = 1 - F(x)$, λ denote the arrival rate, and $\rho = \lambda E[X]$ be the load. Define $\rho(x) = \lambda \int_0^x t dF(t)$. Here $\rho(x)$ can be interpreted as the load made up by jobs with size $< x$. Then the conditional mean response time for a job of size x, $E[T(x)]$, under $M/GI/1/SRPT$ was first derived by Schrage & Miller [19] and is equal to

$$E[T(x)] = \int_0^x \frac{dt}{1-\rho(t)} + \frac{\lambda x^2 \bar{F}(x)}{2(1-\rho(x))^2} + \frac{\lambda \int_0^x t^2 dF(t)}{2(1-\rho(x))^2}, \quad (1)$$

with $E[T] = E[E[T(x)]]$. Despite the existence of this result, due to its complex form, understanding the behavior of $E[T]$ under SRPT is difficult. For example, it is hard to determine the impact of job size variability and load on this formula. Further, calculating $E[T]$ numerically is nontrivial. The goal of this paper is to provide insight into the behavior of $E[T]$ by studying SRPT in heavy traffic. Our main results are described in asymptotic notation. We write $f(x) \sim g(x)$ as $x \to a$ iff $\lim_{x\to a} |f(x)/g(x)| = 1$. Similarly, $f(x) = o(g(x))$ denotes $\lim_{x\to a} |f(x)/g(x)| = 0$, and $f(x) = \Theta(g(x))$ denotes $0 < \liminf_{x\to a} |f(x)/g(x)| \le \limsup_{x\to a} |f(x)/g(x)| < \infty$.

There is only a limited amount of prior work studying the heavy-traffic behavior of SRPT. Recently, Bansal [2] characterized the heavy-traffic behavior of SRPT in the M/M/1. Soon after, Bansal & Gamarnik [3] studied the mean response time of SRPT in an M/GI/1 queue with a Pareto job size distribution. Pechinkin [15] studies the heavy-traffic queue length distribution under SRPT given Pareto job sizes and rapidly varying job size distributions and characterizes the distributional limit in each case. Pechinkin does not explicitly consider the growth rate of $E[T]$, but does provide results for the case of Pareto job sizes in the text. Finally, Down et al.[6] derives a fluid limit for the conditional response time for a job of size x under job size distributions with finite support.

Our work extends [2, 15, 3] to the general M/GI/1 setting and provides an explicit characterization of the impact of the job size distribution on the heavy-traffic growth rate. Our results show that the heavy-traffic behavior of $E[T]$ under SRPT depends on a measure $G(x) = \rho(x)/\rho$ and the Matuszewska index[14] of the job size distribution.

Definition 1. *Let $f(\cdot)$ be positive,*

- *The upper Matuszewska index $\alpha(f)$ is the infimum of α for which, for some $C = C(\alpha) > 0$ and all $\Lambda > 1$,*

$$\frac{f(\lambda x)}{f(x)} \le C\{1+o(1)\}\lambda^{\alpha} \ (x \to \infty) \text{ uniformly in } \lambda \in [1, \Lambda];$$

- *The lower Matuszewska index $\beta(f)$ is the supremum of β for which, for some $D = D(\beta) > 0$ and all $\Lambda > 1$,*

$$\frac{f(\lambda x)}{f(x)} \ge D\{1+o(1)\}\lambda^{\beta} \ (x \to \infty) \text{ uniformly in } \lambda \in [1, \Lambda];$$

3. MAIN RESULTS AND DISCUSSION

In this section we present our main results characterizing the heavy-traffic growth rate of $E[T]$ under SRPT and some numeric results as well. We omit all proofs, which can be found in [13].

3.1 Heavy-traffic results

The first theorem we present is a comparison of the heavy-traffic behavior of SRPT to the heavy-traffic behavior of other common policies – Processor Sharing (PS) and First Come First Served (FCFS). Recall that, in an $M/G/1$ system, the mean response time under PS (FCFS) is $E[T] = \Theta(\frac{1}{1-\rho})$ regardless of the job size distribution as long as it has finite first (second) moment [12]. The following theorem shows that the growth rate of $E[T]$ for SRPT can be slower than that of PS and FCFS as $\rho \to 1$. Further, it shows that the growth rate depends on the job size distribution.

Theorem 2. *In an $M/GI/1$ SRPT queue, as $\rho \to 1$,*

$$E[T] = \begin{cases} \Theta\left(\frac{1}{1-\rho}\right) & F(x) \text{ has bounded support} \\ o\left(\frac{1}{1-\rho}\right) & \text{Otherwise} \end{cases}.$$

Theorem 2 shows that not only does SRPT minimize $E[T]$, it also provides an improvement which is larger than a constant factor if and only if the job size distribution is unbounded. The following theorem will characterize the growth rate of $E[T]$ explicitly:

Theorem 3. *In an $M/GI/1$ SRPT queue, if $F(x)$ has unbounded support, then as $\rho \to 1$,*

$$E[T] = \begin{cases} \Theta\left(\frac{1}{(1-\rho)G^{-1}(\rho)}\right) & \alpha(\bar{F}) < -2 \\ \Theta\left(\log\left(\frac{1}{1-\rho}\right)\right) & \beta(\bar{F}) > -2 \end{cases}.$$

Theorem 3 illustrates the precise impact of the job size distribution on the growth rate of $E[T]$. It turns out that the growth rate is determined by $G(\cdot)$ and the Matuszewska index. Both of which are related to the tail of the job size distribution. It shows that job size distributions with heavier tails have $E[T]$ that increases more slowly as $\rho \to 1$. And, once the tail is heavy enough, the growth rate becomes $\Theta\left(\log\left(\frac{1}{1-\rho}\right)\right)$, which is "independent" of the distribution (though the constant may be different). Notice that the exact constant for $E[T]$ depends on the distribution. Thus, one cannot hope to provide a constant without using explicit information about the job size distribution. However, given special classes of distribution, it is possible to get the exact constant for $E[T]$. To illustrate this fact, the following theorem characterizes the asymptotic mean response time for regularly varying job size distributions (RV_α), which have $\bar{F}(x) = L(x)x^{\alpha}$ where $L(\cdot)$ is a slowly varying function (i.e. $L(ax)/L(x) \to 1$ as $x \to \infty$ for every $a > 0$).

Theorem 4. *In an $M/GI/1$ SRPT queue, if $\bar{F}(x) \sim RV_\alpha$, then as $\rho \to 1$,*

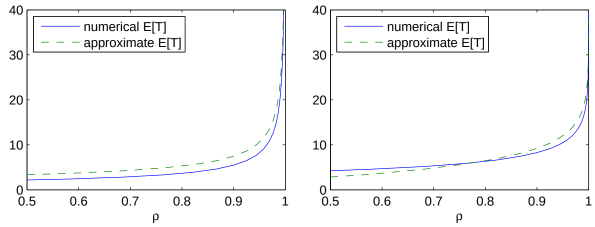
$$E[T] \sim \begin{cases} \frac{(\pi/(\alpha+1))}{2\sin(\pi/(\alpha+1))} \cdot \frac{E[X^2]}{(1-\rho)G^{-1}(\rho)} & \alpha < -2 \\ \frac{1}{(-\alpha)(\alpha+2)} \cdot E[X]\log\left(\frac{1}{1-\rho}\right) & \alpha > -2 \end{cases}.$$

Note that for a regularly varying distribution $\bar{F}(x) \sim RV_\alpha$, we have $\alpha(\bar{F}) = \beta(\bar{F}) = \alpha$ and so Theorem 3 is still applicable. However, Theorem 4 provides a refined result. This theorem generalizes recent results from Bansal & Gamarnik [3] and Pechinkin [15], who each analyze the Pareto distribution, which is a special case of regularly varying distributions. Additionally, note that this result is valid for tails that are of rapid variation (corresponding to $\alpha = -\infty$), which satisfy $\bar{F}(xy)/\bar{F}(x) \to 0$ as $x \to \infty$ for $y > 1$. Most light-tailed distributions fit into this framework.

3.2 Beyond heavy-traffic

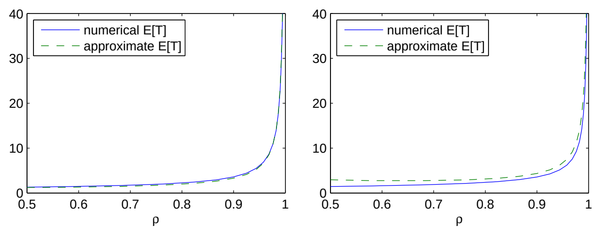
The theorems that have just been presented characterize the heavy-traffic performance of SRPT, however, heavy-traffic limits often provide simple descriptions of queueing models that can be used to develop approximations outside of heavy-traffic, e.g., [20, 10, 8]. In this section, we will illustrate that Theorems 3 and 4 can also be used in this manner. In particular, we will use numeric experiments to illustrate that the form of $E[T]$ provides an accurate approximation of $E[T]$ outside of heavy-traffic.

We will focus our numeric experiments on the cases of the Pareto and Weibull job size distributions. The Pareto



(a) $\alpha = 3$ (b) $\alpha = 1.5$

Figure 1: Comparison of approximate $E[T]$ with numeric $E[T]$ as a function of ρ in the case of Pareto job sizes outside of heavy-traffic.



(a) $\alpha = 2$ (b) $\alpha = 1$

Figure 2: Comparison of approximate $E[T]$ with numeric $E[T]$ as a function of ρ in the case of Weibull job sizes outside of heavy-traffic.

distribution is probably the most popular heavy-tailed distribution, and is widely used as a model for the tails of real world workloads. The Weibull distribution is another common job size distribution since it can mimic the behavior of many other distributions with increasing or decreasing failure rate.

The first example we consider is the case of Pareto job sizes. For $X \sim Pareto(\alpha) \in RV_{-\alpha}$, we have $G^{-1}(\rho) = x_m(1-\rho)^{\frac{1}{1-\alpha}}$, and we get the asymptotic mean response time for Pareto job sizes immediately from Theorem 4. Now we use this asymptotic formula as an approximation for $E[T]$ and compare it to numerical calculation of the exact mean response time from (1). For Pareto job sizes with various parameters: $\alpha = 3, 1.5$, it shows the approximate $E[T]$ and the numerical $E[T]$ as a function of ρ for a wider range of loads, which illustrates that the asymptotic formula is a good approximation even outside of heavy-traffic.

The second example we consider is the case of Weibull job sizes. Note that the $M/M/1$ is a special case of Weibull obtained by setting $\alpha = 1$. Notice that $X \sim Weibull(\alpha) \in RV_{-\infty}$. So we can calculate $G^{-1}(\rho)$ and apply Theorem 4 to attain the asymptotic results for Weibull job sizes:

$$E[T] \sim \frac{E[X^2]}{2} \frac{1}{(1-\rho)\cdot \mu^{-1/\alpha} \log(\frac{1}{1-\rho})^{1/\alpha}} \quad as\ \rho \to 1 \quad (2)$$

Again, we now use the above asymptotic formula as an approximation for $E[T]$ and compare it to numerical calculation of the exact mean response time from (1). The numerical $E[T]$ and the approximate $E[T]$ for Weibull distributions with various parameters ($\alpha = 1, 2$) are shown in Figure 2. Again, the approximation seems accurate even when ρ is not very large.

4. REFERENCES

[1] L. Andrew, M. Lin, and A. Wierman. Optimality, fairness, and robustness in speed scaling designs. In *Proc. of ACM Sigmetrics*, 2010.

[2] N. Bansal. On the average sojourn time under M/M/1/SRPT. *Oper. Res. Letters*, 22(2):195–200, 2005.

[3] N. Bansal and D. Gamarnik. Handling load with less stress. *Queueing Systems*, 54(1):45–54, 2006.

[4] O. Boxma and B. Zwart. Tails in scheduling. *Perf. Eval. Rev.*, 34(4):13–20, 2007.

[5] H.-L. Chen, J. Marden, and A. Wierman. On the impact of heterogeneity and back-end scheduling in load balancing designs, 2009.

[6] D. Down, H. C. Gromoll, and A. L. Puha. The fluid limit of an SRPT queue. Under submission.

[7] C. Goerg. Evaluation of the optimal SRPT strategy with overhead. *IEEE Trans. on Comm.*, 34(4):338–344, 1986.

[8] P. Hall. Heavy traffic approximations for busy period in an $M/G/\infty$ queue. *Stochastic Processes and their Applications*, 19(2):259 – 269, 1985.

[9] M. Harchol-Balter, B. Schroeder, M. Agrawal, and N. Bansal. Size-based scheduling to improve web performance. *ACM Transactions on Computer Systems*, 21(2), 2003.

[10] J. M. Harrison. The heavy traffic approximation for single server queues in series. *Journal of Applied Probability*, 10(3):613–629, 1973.

[11] M. Hu, J. Zhang, and J. Sadowsky. A size-aided opportunistic scheduling scheme in wireless networks. In *Globecom*, 2003.

[12] L. Kleinrock. *Queueing Systems*, volume I. Theory. John Wiley & Sons, 1975.

[13] M. Lin, A. Wierman, and B. Zwart. Heavy-traffic analysis of mean response time under shortest remaining processing time. *Under submission*.

[14] C. M. G. N. H. Bingham and J. L. Teugels. Regular variation. *Cambridge University Press*, 1987.

[15] A. V. Pechinkin. Heavy traffic in a system with a discipline of priority servicing for the job of shortest remaining length with interruption. *Mat. Issled.*, (89, Veroyatn. Anal.):85–93, 97, 1986.

[16] A. V. Pechinkin. The MAP/G/1 queue with SRPT service discipline. *Theory of Prob. and its App.*, 45(3):532–539, 2000.

[17] I. A. Rai, G. Urvoy-Keller, M. Vernon, and E. W. Biersack. Performance modeling of FB based scheduling in packet switched networks. In *Proc. of ACM Sigmetrics-Performance*, 2004.

[18] L. E. Schrage. A proof of the optimality of the shortest remaining processing time discipline. *Operations Research*, 16:678–690, 1968.

[19] L. E. Schrage and L. W. Miller. The queue M/G/1 with the shortest remaining processing time discipline. *Operations Research*, 14:670–684, 1966.

[20] W. Whitt. Efficiency-driven heavy-traffic approximations for many-server queues with abandonments. *Management Science*, 50(10):1449–1461, 2004.

[21] A. Wierman. Fairness and classifications. *Perf. Eval. Rev.*, 34(4):4–12, 2007.