INFORMS 2018 Open Problem Contribution

Isaac Grosof*

March 12, 2019

1 M/G/k/SRPT Under Medium Load

The Shortest-Remaining-Processing-Time (SRPT) scheduling policy has been deployed in many computer systems, such as web servers [7], networks [9], databases [4] and operating systems [1]. SRPT has also received extensive study by queueing theorists. In 1966, the mean response time for SRPT was first derived [11], and in 1968 SRPT was shown to minimize mean response time [10]. However, these results are only known for *single-server* systems. Much less is known for *multiserver* systems, such as the M/G/k.

The first analytic results on the M/G/k/SRPT were given by Grosof, Scully and Harchol-Balter [3] in 2018. They proved analytic bounds on the mean response time of the M/G/k/SRPT. They showed these bounds are tight in the heavy-traffic limit, namely as load approaches capacity. However, their bounds leave open the question of mean response time of under *medium* load.

The medium-load regime has deep importance to the design of computer systems. When designing a multiserver computer system, we want to achieve high utilization while also keeping mean response time low. On a graph of mean response time $E[T]$ by load ρ, the region of importance is the one known as the "knee" of the curve. The knee refers to the load where mean response time begins to increase above its minimum, the mean job size. This motivates the following question:

> In an M/G/k with SRPT scheduling, where is the knee of the mean response time curve?

To formalize this question, we pose the following open problem:

For a given $\epsilon > 0$, what is the maximum load $\rho_H(\epsilon)$ such that

$$E[T]^{\text{M/G/k/SRPT}} \leq (1+\epsilon)E[S], \tag{1}$$

where T denotes the response time distribution and S denotes the job size distribution? We are particularly interested in this question for ϵ in the range 0.1 to 1.

1.1 Prior Work: M/G/k/FCFS

Note that the equivalent question for a First-Come-First-Served system, the M/G/k/FCFS, is far better understood. Early analysis by Kingman [8], and later by Daley [2] bounded mean response time based on the first two moments of the job size distribution. Osogami et al. [6] exactly analyzed mean response time in a M/Ph/k/FCFS system. Since an arbitrary distribution can be approximated by a phase-type distribution, this analysis allows an approximation of mean response time in the M/G/k/FCFS. Most recently, Gupta and Osogami [5] showed how to use the entire sequence of moments of a job size distribution to give tight bounds on mean response time. These results allow us to tightly bound mean response time in the M/G/k/FCFS at all loads, and answer questions like (1) for the M/G/k/FCFS.

1.2 Simulation

While the M/G/k/FCFS system is better understood than the M/G/k/SRPT, simulations indicate that the SRPT system offers superior performance under medium load. SRPT's excellent empirical performance in this regime makes theoretical analysis especially important.

*igrosof@cs.cmu.edu. Carnegie Mellon University, Computer Science Department, Pittsburgh, PA, USA

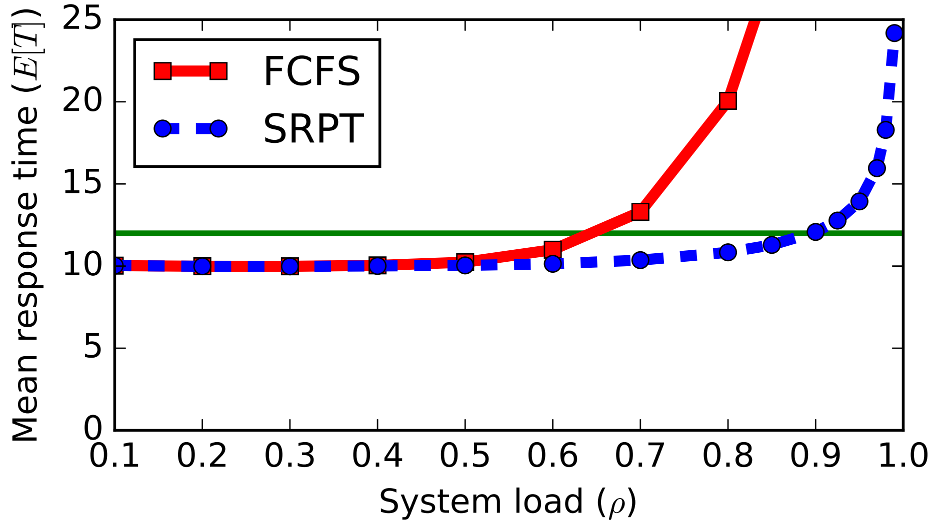

Figure 1: Mean response time (E[T]) versus load, under SRPT and FCFS scheduling. Load at which E[T] first exceeds 1.2 times mean job size E[S] is about 0.65 under FCFS and about 0.9 under SRPT. We use $k = 10$ servers. The job size distribution S is hyperexponential with $E[S] = 10$ and $C^2 = 10$.

Figure 1 shows mean response time under FCFS and SRPT scheduling policies for a M/G/k system. The job size distribution is hyperexponential with mean 10 and $C^2 = 10$, and there are $k = 10$ servers in the system. Under FCFS scheduling the load at which mean response time first exceeds mean job size by 20%, $\rho_H(0.2)$, is approximately 0.65. In contrast, under SRPT scheduling $\rho_H(0.2)$ is approximately 0.9. In this scenario, we see that SRPT scheduling allows much better server utilization while maintaining the same mean response time.

While we empirically observe that SRPT scheduling yields relatively low mean response time under medium load in the M/G/k, we have no theoretical justification for this behavior. In particular, it remains to be shown whether this behavior persists across all job size distributions and any amount of servers.

1.3 Conclusion

We propose an important open problem: analyzing the mean response time in the M/G/k/SRPT under medium load. While mean response time in the M/G/k/SRPT has recently been tightly bounded under heavy traffic, little to nothing is known under medium load. The excellent empirical mean response time of the M/G/k/SRPT under medium load motivates its further study.

References

[1] Bunt, R. B. Scheduling techniques for operating systems. *Computer 9*, 10 (1976), 10–17.

[2] Daley, D. J. Some results for the mean waiting-time and workload in gi/gi/k queues. *Frontiers in queueing: models and applications in science and engineering* (1997), 35–59.

[3] Grosof, I., Scully, Z., and Harchol-Balter, M. Srpt for multiserver systems. *Performance Evaluation 127-128* (2018), 154 – 175.

[4] Guirguis, S., Sharaf, M. A., Chrysanthis, P. K., Labrinidis, A., and Pruhs, K. Adaptive scheduling of web transactions. In *Data Engineering, 2009. ICDE'09. IEEE 25th International Conference on* (2009), IEEE, pp. 357–368.

[5] Gupta, V., and Osogami, T. On markov–krein characterization of the mean waiting time in m/g/k and other queueing systems. *Queueing Systems 68*, 3 (Jul 2011), 339.

[6] Harchol-Balter, M., Osogami, T., Scheller-Wolf, A., and Wierman, A. Multi-server queueing systems with multiple priority classes. *Queueing Systems 51*, 3 (Dec 2005), 331–360.

[7] Harchol-Balter, M., Schroeder, B., Bansal, N., and Agrawal, M. Size-based scheduling to improve web performance. *ACM Trans. Comput. Syst. 21*, 2 (May 2003), 207–233.

[8] Kingman, J. F. C. Inequalities in the theory of queues. *Journal of the Royal Statistical Society: Series B (Methodological) 32*, 1 (1970), 102–110.

[9] Montazeri, B., Li, Y., Alizadeh, M., and Ousterhout, J. Homa: A receiver-driven low-latency transport protocol using network priorities. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication* (New York, NY, USA, 2018), SIGCOMM '18, ACM, pp. 221–235.

[10] Schrage, L. Letter to the editor-a proof of the optimality of the shortest remaining processing time discipline. *Operations Research 16*, 3 (1968), 687–690.

[11] Schrage, L. E., and Miller, L. W. The queue M/G/1 with the shortest remaining processing time discipline. *Operations Research 14*, 4 (1966), 670–684.