Nudge: Stochastically Improving upon FCFS

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

Kunhe Yang
yangkunhe19@gmail.com
Tsinghua University
Institute for Interdisciplinary Information Sciences
Beijing, China

Ziv Scully
zscully@cs.cmu.edu
Carnegie Mellon University
Computer Science Department
Pittsburgh, PA, USA

Mor Harchol-Balter
harchol@cs.cmu.edu
Carnegie Mellon University
Computer Science Department
Pittsburgh, PA, USA

ABSTRACT

The First-Come First-Served (FCFS) scheduling policy is the most popular scheduling algorithm used in practice. Furthermore, its usage is theoretically validated: for light-tailed job size distributions, FCFS has weakly optimal asymptotic tail of response time. But what if we don't just care about the asymptotic tail? What if we also care about the 99th percentile of response time, or the fraction of jobs that complete in under one second? Is FCFS still best? Outside of the asymptotic regime, only loose bounds on the tail of FCFS are known, and optimality is completely open.

In this paper, we introduce a new policy, Nudge, which is the first policy to provably stochastically improve upon FCFS. We prove that Nudge simultaneously improves upon FCFS at *every* point along the tail, for light-tailed job size distributions. As a result, Nudge outperforms FCFS for every moment and every percentile of response time. Moreover, Nudge provides a multiplicative improvement over FCFS in the asymptotic tail. This resolves a long-standing open problem by showing that, counter to previous conjecture, FCFS is not strongly asymptotically optimal.

This paper represents an abridged version of [2].

ACM Reference Format:
Isaac Grosof, Kunhe Yang, Ziv Scully, and Mor Harchol-Balter. 2021. Nudge: Stochastically Improving upon FCFS. In *Abstract Proceedings of the 2021 ACM SIGMETRICS / International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '21 Abstracts), June 14–18, 2021, Virtual Event, China.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3410220.3460102

1 INTRODUCTION

While advanced scheduling algorithms are a popular topic in theory papers, it is unequivocal that the most popular scheduling policy used in practice is still First-Come First-Served (FCFS). There are many reasons for the popularity of FCFS. For instance, FCFS is easy to implement, and has a feeling of being fair.

However, there are also theoretical arguments for why one should use FCFS. For one thing, FCFS minimizes the *maximum* response time for *any* finite arrival sequence of jobs. By *response time* we mean the time from a job's arrival to its completion.

For another thing, in an M/G/1 with a light-tailed job size distribution, FCFS is known to have a *weakly optimal* asymptotic tail of response time [1]. Specifically, using T to denote response time, the asymptotic tail under FCFS is of the form $\mathbf{P}\{T^{\text{FCFS}} > t\} \sim C_{\text{FCFS}} e^{-\theta^* t}$, where θ^* is known to be optimal, while the optimality of C_{FCFS} is an open problem [1]. The asymptotic tail growth under FCFS has been compared with more sophisticated policies [1]. It has been shown that, for light-tailed job size distributions, the tail of response time under Processor-Sharing, Preemptive Last-Come-First-Served, and Shortest-Remaining-Processing-Time (SRPT) each take the asymptotic form of $\mathbf{P}\{T > t\} \sim C' e^{-\theta' t}$, where θ' is the *worst possible* exponential decay rate over all work-conserving scheduling policies. Roughly, FCFS's tail exponent θ^* arises from the tail of the workload distribution, while the other policies' tail exponent θ' arises from the tail of the busy period distribution, which is much larger under light-tailed job size distributions.

In this paper, we choose to focus on the case of light-tailed job size distributions. Light-tailed job size distributions show up naturally in workloads where all the transactions are of the same type (say shopping). Also, many natural distributions, like the Normal distribution, Exponential distribution, and all Phase-type distributions, are light-tailed. Finally, while heavy-tailed job size distributions are certainly prevalent in empirical workloads, in practice, these heavy-tailed workloads are often *truncated*, which immediately makes them light-tailed.

Within the world of light-tailed job size distributions, FCFS is viewed as the best policy. However, while FCFS has a weakly optimal *asymptotic* tail, it is not best at minimizing $\mathbf{P}\{T > t\}$ for *all* t. In practice, one cares less about the asymptotic case than about particular t. For example, one might want to minimize the fraction of response times that exceed $t = 0.5$ seconds, because such response times are noticeable by users. One might also want to meet several additional Service Level Objectives (SLOs) where one is charged for exceeding particular response time values, such as $t = 1$ minute, or $t = 1$ hour. SLOs are very common in the computing literature, in service industries, and in healthcare. Unfortunately, different applications have different SLOs. This leads us to ask:

SIGMETRICS '21 Abstracts, June 14–18, 2021, Virtual Event, China

ACM ISBN 978-1-4503-8072-0/21/06.
https://doi.org/10.1145/3410220.3460102

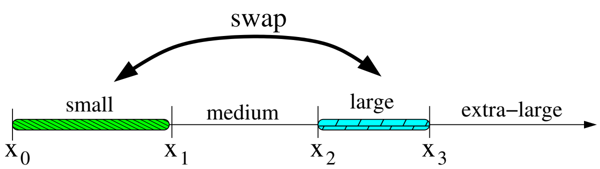

Figure 1: The Nudge algorithm.

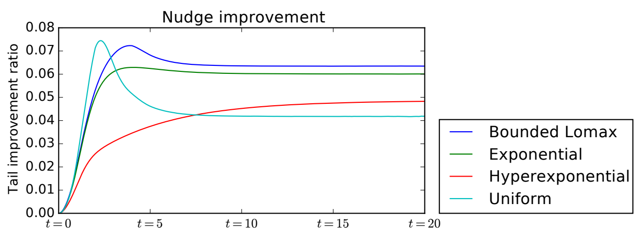

Figure 2: Empirical tail improvement of Nudge over FCFS in an M/G/1. Tail improvement ratio $= 1 - \mathbf{P}\{T^{\text{Nudge}} > t\}/\mathbf{P}\{T^{\text{FCFS}} > t\}$. **Job size distributions,: Uniform**(0,2)**; Exp**(1)**; Hyperexponential**($\mu_1 = 2, p_1 = 0.8, \mu_2 = 1/3, p_2 = 0.2$)**; BoundedLomax**($\lambda = 2$, max $= 4$, $\alpha = 2$)**. Distributions in legend ordered by asymptotic improvement. Simulations of** 10 **billion arrivals. Load** $\rho = 0.8$. **Nudge parameters:** $x_1 = 1, x_2 = 1, x_3 = \infty$.

> *When considering* $\mathbf{P}\{T > t\}$*, is it possible to strictly improve upon FCFS for* <u>*all*</u> *values of* t*?*

We are motivated by the fact that, for lower values of t, SRPT is better than FCFS, although FCFS clearly beats SRPT for higher values of t, as FCFS is weakly asymptotically optimal while SRPT is asymptotically pessimal. SRPT also minimizes mean response time, which is closely related to lower values of t. This motivates us to consider whether prioritizing small jobs might have some benefit, even in the world of light-tailed job size distributions.

We ask more specifically:

> *Can partial prioritization of small jobs lead to a strict improvement over FCFS? Specifically, is there a scheduling policy which strictly improves upon FCFS with respect to* $\mathbf{P}\{T > t\}$*, for* every *possible* t *including large* t*?*

This paper answers the above question in the affirmative. We will define a policy, which we call *Nudge*, whose response time tail is *provably* better than that of FCFS for *every* value of t, assuming a light-tailed job size distribution[1]. We say that Nudge's response time *stochastically improves upon* that of FCFS, in the sense of stochastic dominance. Moreover, we prove that the asymptotic tail of response time of Nudge is of the form $P\{T^{\text{Nudge}} > t\} \sim C_{\text{Nudge}} e^{-\theta^* t}$, with optimal decay rate θ^* and a superior leading constant $C_{\text{Nudge}} < C_{\text{FCFS}}$. Thus, we demonstrate that FCFS is *not* strongly optimal, answering an open problem posed by Boxma and Zwart [1]. In particular, this disproves a conjecture of Wierman and Zwart [4].

The intuition behind the Nudge algorithm is that we'd like to basically stick to FCFS, which we know is great for handling the extreme tail (high t), while at the same time incorporating a little bit of prioritization of small jobs, which we know can be helpful for the mean and lower t. We need to be careful, however, not to make too much use of size, because Nudge still needs to beat FCFS for high t; hence we want just a little "nudge" towards prioritizing smalls.

We now describe the Nudge algorithm. Imagine that the job size distribution is divided into size regions, as shown in Fig. 1, consisting of small, medium, large, and extra large jobs. Most of the time, Nudge defaults to FCFS. However, when a "small" job arrives and finds a "large" job immediately ahead of it in the queue, then we swap the positions of the small and large job in the queue. The one caveat is that a job which has already swapped is ineligible for further swaps. The size cutoffs defining small and large jobs will be defined later in this paper.

The degree of the tail improvement of Nudge over FCFS is non-trivial. In Fig. 2, we see that for many common light-tailed job size distributions, Nudge results in a multiplicative improvement of 4-7% throughout the tail. In this paper, we show that with low load and a high-variability job size distribution, Nudge's improvement can be as much as 10-15% throughout the tail. Furthermore, the fact that Nudge stochastically improves upon FCFS means that it beats FCFS for all moments of response time, all percentiles of response time, and all combinations thereof.

We conclude this paper by presenting an exact analysis of the performance of Nudge. Nudge does not fit into any existing framework for M/G/1 transform analysis, including the recently developed SOAP framework [3]. Nonetheless, we give a tagged-job analysis of Nudge, deriving the Laplace-Stieltjes transform of response time of Nudge.

Our paper makes the following contributions (see [2]):

- We introduce the Nudge policy.
- We prove that with appropriately chosen parameters, Nudge stochastically improves upon FCFS for light-tailed[1] job size distributions; we also give a simple expression for such parameters. Moreover, we prove that Nudge achieves a multiplicative asymptotic improvement over FCFS.
- We derive the Laplace-Stieltjes transform of response time under Nudge.
- We empirically demonstrate the magnitude of Nudge's stochastic improvement over FCFS. We also discuss how to tune Nudge's parameters for best performance.

ACKNOWLEDGEMENTS

This research was supported by NSF-CMMI-1938909, NSF-CSR-1763701, and a Google 2020 Faculty Research Award.

REFERENCES

[1] Onno Boxma and Bert Zwart. 2007. Tails in Scheduling. *ACM SIGMETRICS Performance Evaluation Review* 34, 4 (March 2007), 13–20.
[2] Isaac Grosof, Kunhe Yang, Ziv Scully, and Mor Harchol-Balter. 2021. Nudge: Stochastically Improving upon FCFS. *Proc. ACM Meas. Anal. Comput. Syst.* 5, 2 (June 2021), 28 pages.
[3] Ziv Scully, Mor Harchol-Balter, and Alan Scheller-Wolf. 2018. SOAP: One Clean Analysis of All Age-Based Scheduling Policies. *Proc. ACM Meas. Anal. Comput. Syst.* 2, 1, Article 16 (April 2018), 30 pages.
[4] Adam Wierman and Bert Zwart. 2012. Is Tail-Optimal Scheduling Possible? *Operations Research* 60, 5 (Oct. 2012), 1249–1257.

[1] Technically, a Class I job size distribution. See [2].