

Comments on the revised version

March 8, 2018

Let us first sincerely thank the reviewers for their relevant and in-depth remarks once again. Their comments helped us further the discussion of the results, and find a typographical error in the presentation of the results that passed our scrutiny.

We detail our answers to reviewer's comments (bold).

1 Reviewer 1

I basically liked this paper and support its publication. I have only very few and minor comments.

In sect 1.1 you claim 2 "new" online approaches, and in 1.2 you call them "original". Your work is valid even if you admit that both have been suggested before.

Last paragraph of 1.2 is the third repetition of this. It is redundant, despite the fact that some reviewer didn't get it.

You are right. We rephrased this in a more neutral way.

In sect 5.1 you explain that you set $\lambda=1$ because due to the resampling the distributions are stationary. That is correct. But it would have been good to also conduct simulations on the original traces, which are probably not stationary, and therefore there is more to gain from switching between policies. Also, instead of starting the simulation of each period from an empty system, why not start from the situation at the end of the previous period?

Regarding your first point, good remark, we based this experimental design around previous results from "Tuning Backfilling Queues" (JSSPP17) that show some degree of stationarity using a train/test procedure. We made this clearer in the text. We indeed expect that non-stationary policy adaptation may allow more improvements (while simultaneously being harder to achieve).

Your second point is also relevant, yes that should improve the results. But we can see that the 'full feedback' simulated policy here is almost as good as the best policy in the set, so we did not worry about this. In a non-stationary case, it may indeed be necessary to properly take care of this.

Typo at end of 6.1.3: sentence with [24] appears twice.

Right, we fixed this.

2 Reviewer 2

In comparison to the first version, the authors have made several small changes that results in minor improvements of the manuscript. Nevertheless, the main problems of the manuscript remain. Therefore, the position of the reviewer remains unchanged.

The authors present a study based on a selection of several strategies and parameters. They execute several simulations and show that using a feedback mechanism improvements over the standard approach without feedback are possible. This results is not surprising considering previous results some of which are referenced within the manuscript. It must be noted that

the authors do not suggest a specific new scheduling algorithm and show that this algorithm results in significant improvements over existing methods.

We regret that the detailed answer to your comments did not convince you. We disagree with this statement. We indeed provide a new scheduling algorithm, which adapts to the situation by selecting the correct reordering policy. This algorithm results in significant improvements over existing methods. We thank you for your in-depth remarks once again. We integrated these comments into the minor revision.

The authors mention "reducing uncertainty" in Section 1.1. It is not clear why this reduction is relevant within the manuscript (except as a justification for adding noise?) The first paragraph of Section 1.1 is relevant to the content of the manuscript. The study of the manuscript uses old traces that are far from extreme scale parallel computing. The relevance to distributed computing is not clear since only traces from parallel computers are used.

Why do the authors discuss conservative backfilling in Section 2 although it does not matter in the rest of the manuscript? What is the meaning of "dependency of scheduling metrics on the workload"? The authors only apply a single scheduling metric (average waiting time) and do not discuss such dependency. Why is it important that previous work does not address this dependency in the context of this manuscript?

The authors mention duality between cumulative and maximal scheduling costs. Why is this duality relevant in the context of this manuscript? In Section 3.3 the authors claim that "Section 4.2 will outline our approach to address the biobjective aspect of this problem". However, Section 4.2. is very brief and does not discuss such aspect. The reviewer did not find any discussion in the rest of the manuscript. Altogether not all references cited in the manuscript are relevant for the content of the manuscript.

The design of the study is contradictory: on the one hand, the authors only select a single metric since this metric "is one of the more commonly used objectives". On the other hand, they select 12 reordering strategies without giving any justification based on the strategy. They simply state that they want the search space "being as semantically diverse as possible" without stating that they achieved this goal with the selection of these strategies. Why do they need all the 12 strategies and why not using more strategies? Again they only remark that the strategies "include most policies from related works that we are aware of." Such justification is hardly sufficient for a scientific study.

In Section 5, the authors claim that "the distribution of the job submission process does not radically differ between consecutive periods." This claim is not true for every period length. For instance, when using 12 hours as period length, two consecutive periods may cover day and night with different submission processes.

The allocation of jobs to periods based on the submission time may be problematic. For instance a long running job may be submitted in one period, waiting for execution in the next period and then executing in the next several periods. Therefore, it influences several periods. The authors briefly mention such dependencies (Section 5.3) and state that there may be a bias without discussing this issue.

The authors still keep the decaying discount factor λ without using it in the study. In an exploratory study, the authors can either ignore such factor because it is not relevant (and saying why this claim holds) or they must include it into the study.

The authors introduce a noise factor to deal with imprecision. This way they introduce an approach to model imprecision. But they do not show that their model is suitable in reality.

In Section 5.2, the authors claim that the length of the period has little impact here as the cost metric corresponds to the cumulative simulated waiting time for each policy. As already mentioned, this may not be true for periods that are small compared to the waiting time of jobs.

Figure 3 does not illustrate the process of the ϵ -greedy strategy since there is no ϵ in Figure 3.

Section 6.1.1 mentions cores in traces although several of the old traces use machines that do not have multicore processors.

The authors use 40 hours as threshold and justify this selection by referring to another publication. The reader does not know about the arguments without reading this other publication. Since the design of the study is at the core of this manuscript, such justification is not acceptable.

The authors state that they have simulated a total of 92400 years. However, this absolute figure is not relevant without showing the coverage of the problem space.

There are only few language problems in the manuscript: Software and work are used in the singular form. The authors must adjust the verb correspondingly (Section 1.1) workload in Fig. 2 and 3

3 Reviewer 3

I think that the simulation model presented in this paper is carefully created and it shows convincing results. However, there are two questions about the discussion for the simulation results:

The results in Table 3 and Table 4 show that the performance of the static strategy, LQF, is comparable with the presented online tuning strategies, Full and Bandit. Which strategy is better in the real world? We can say that LQF may be preferable because they are less complicated and run with smaller scheduling overhead. The authors should add the discussion to compare LQF and the online tuning strategies.

First of all, we corrected a crucial typographical error in the text that slipped our scrutiny at the time of running the build chain for the publication. The S and L were inverted in both Table 3 and Figure 4. This resolves a blatant consistency issue with other publications, most crucially "Tuning Backfilling Queues" (JSSPP17).

You are right, the discussion you suggest is warranted. We added it to paragraph **SQF versus online tuning** in Section 6.3.2:

Table 3 shows that the SQF policy works quite well on the traces considered here on average, and under this resampling scheme. The variability however, is high, and this policy is not the best on all traces. It is interesting to have an experimental evaluation of exactly how much one gains on average by using an online policy versus a leave-one-trace-out majority vote policy choice. However, a good estimation of this quantity is tricky. Indeed, we expect the change in the relative performance of policies to increase with two factors. First, simply adding workload traces to our selection will naturally show more cases where behavior is different. Second, factoring in the topology model of the machines should increase the heterogeneity between platforms and therefore increase the dependence of the simulation performance on the trace/platform. Such a study would be much more tedious. Here the need for adaptivity is actually understated by the simulation approach.

The authors conclude that simulation-based strategy is best but the bandit-based strategy is easier to use and cheaper to run. How easy and how cheaper can we run the bandit-based strategy? The authors should add the quantitative discussion how the bandit-based strategy is easy to use and saves cost (or scheduling overhead) compared with the simulation-based strategies.

Right, we need to make this clearer. We added this discussion to paragraph **SQF versus online tuning** in Section 6.3.2:

While we can not conclude on how much one gains by not using SQF in this case, we can comment on the added complexity and computing overhead of using the simulation-based and bandit online

strategies. In both these cases, the scheduling decisions do not have an overhead, as the policy is fixed during the period (one day, or week). The bandit strategy has a negligible computational overhead at period end. Its two computational operations are maintaining a sequential average and generating a pseudo-random number. The simulation-based strategy requires to simulate the system K times every period end, where K is the number of alternative policies considered (here, $K=12$). Using our simulator, This takes a few seconds. Using a more precise simulator with topological modeling, one can argue that the overhead is still manageable since these simulations can be easily done in parallel. The most important constraint is that these simulations should not take too much time compared to the period size, as to not delay the policy switch decision too much.