



Universiteit  
Leiden  
The Netherlands

## Workload modeling and performance evaluation in parallel systems

Minh, T.N.

### Citation

Minh, T. N. (2012, March 22). *Workload modeling and performance evaluation in parallel systems*. Retrieved from <https://hdl.handle.net/1887/18620>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/18620>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/18620> holds various files of this Leiden University dissertation.

**Author:** Tran, Minh Ngoc

**Title:** Workload modeling and performance evaluation in parallel systems

**Issue Date:** 2012-03-22

---

# Chapter 9

## Conclusions

---

Scheduling plays a significant role in producing good performance for clusters and grids. Smart scheduling policies in these systems are essential to enable efficient resource allocation mechanisms. One of the key factors that have a strong effect on scheduling is the workload. This workload problem is associated with four research topics to obtain an effective scheduler, namely workload characterisation, workload modeling, performance evaluation and prediction, and scheduling design. Workload data collected from real systems are the best source for improving our knowledge about performance issues of clusters and grids. Observed features of these workloads are precious sources of clues, which can be utilized to enhance scheduling. To this end, several long-term parallel and grid workloads have been collected [31, 80] and this thesis used these real workloads in the study of workload characterisation, workload modeling, performance evaluation and prediction. Our research resulted in many workload modeling tools, a performance predictor and several useful clues that are essential to develop efficient cluster and grid schedulers.

First of all, this thesis provided a comprehensive characterisation on real cluster and grid workloads, with an emphasis on several statistical features. It was shown that the features long range dependence and temporal burstiness exist strongly in real system job traffic. The analysis also indicated the common presence of the characteristics temporal locality, cross-correlation and spatial burstiness of runtime and parallelism processes. In particular, an approach of quantifying spatial burstiness was suggested and demonstrated experimentally to work well. This efficient approach was used to quantify spatial burstiness in real world data and the results suggested that spatial burstiness deserves more attention from research community, so that correct workloads are used when evaluating scheduling algorithms. In addition, the well-known Bag-of-Tasks (BoT) behaviour was also analyzed and the result showed that users often submit a large number of their jobs under this behaviour. We identified several pattern structures of BoT arrivals and it turned out that the Generalized Pareto distribution is a good fit for BoT interarrivals. It was also illustrated that BoT arrivals are bursty and can have similar or completely contrary structures as job arrivals, with respect to long range dependence and periodicity. In addition to BoT arrivals, statistics like autocorrelation and cross-correlation were also applied to

BoT sizes, runtimes, parallelisms and estimates for a comprehensive analysis of this behaviour. With the common presence of all the mentioned workload features in real world data and their crucial roles on performance issues of clusters, grids and clouds, we argued that more research on workload modeling and performance evaluation is needed.

Once a particular feature is observed in real workloads, it will be interesting to discover its potential impact on scheduling performance. Hence, it is essential to develop representative workload models that can efficiently capture a particular feature. To this end, we have successfully developed a job arrival model with long range dependence and temporal burstiness. This model is based on the multifractal wavelet model (MWM) [46, 85]. MWM is a good choice when it comes to long-range autocorrelations but it should be adapted before applying to system job traffics because we showed that cluster and grid job arrivals do not only exhibit long range dependence but also show temporal burstiness. Experimental results showed that the developed job arrival model can accurately control the temporal burstiness degree and can produce long range dependence well. Moreover, the synthetic job arrival process generated by the model also fits the marginal distribution nicely.

With respect to the workload attributes runtime and parallelism, a new model was proposed to produce several important features such as temporal locality, spatial burstiness, cross-correlation, etc. Especially, these two workload attributes were not modeled separately but were modeled with the job arrival attribute to capture the well-known Bag-of-Tasks behaviour. Hence, this model can be considered as a comprehensive model since it can generate three workload attributes at the same time, which is necessary for simulation use. With a large number of jobs submitted as part of BoTs, it is clear that the temporal-spatial correlation in parallel system workloads is mainly due to this behaviour. Therefore by capturing BoTs, this model helps to study the impact of the temporal-spatial correlation on scheduling performance. Furthermore, this comprehensive representative workload model can be used in many other research aspects. It first helps to generate realistic workloads for the evaluation of newly designed scheduling algorithms. Secondly, the model can be used to evaluate the impact of individual workload characteristics on scheduling performance. Thirdly, because of the ability to capture many workload characteristics, the model is a useful tool for evaluation studies on the impact of interactions between workload characteristics on the performance of clusters and grids.

After the study of workload characterisation and modeling, performance evaluation was investigated in this thesis. Evaluating the performance impact of workloads on scheduling not only helps to enrich the understanding of a system but also provides several important clues that can improve schedulers. Although many workload features were shown to be present commonly in real world data and they certainly have potential impacts on scheduling, these impacts are hardly described in the literature. Therefore, performance issues of numerous workload characteristics, including long range dependence, temporal burstiness, temporal locality and the cross-correlation between runtimes and parallelisms, were investigated.

---

For long range dependence and temporal burstiness, our evaluation results for a single cluster showed that they have severe performance impacts and can significantly affect a scheduling design. In the scenario of a grid, we indicated that in particular situations, grid and cluster jobs may join together to improve scheduling performance. However in most cases, grid and background cluster jobs do not really affect each other. With regard to the correlation between runtime and parallelism, our experiments showed that a positive correlation has more serious impact on parallel system performance than a negative correlation. However, backfilling policies can help to reduce the impact. Furthermore, scheduling design results may differ significantly under different degrees of correlation. With respect to temporal locality, it decreases the scheduling performance of a parallel system but when the system is overloaded, the feature turns to be useful for non-backfilling policies. Therefore, we conclude that any efficient scheduling design should take long range dependence, temporal burstiness, temporal locality and the cross-correlation between runtime and parallelism into account for a good evaluation. To enable such designs, we introduced, besides the developed comprehensive representative workload model, an efficient procedure that can help to tune the degree of the cross-correlation on demand.

Although we found that a non-backfilling algorithm, that utilizes the temporal locality feature well, can be used when a system is overloaded, we also showed that if the system is underloaded, a backfilling policy that restrains the wait time of a job to a certain threshold is a better choice. Designing a particular backfilling policy is completely feasible because we developed an efficient runtime predictor with high accuracy in this thesis. This predictor takes into account the problems of under/over-estimations, which are important for backfilling because underestimating a job has it killed by the system while overestimating a job gives it less chance to be backfilled.

In addition to the research achievements summarised above, the results of our study have opened several research opportunities for future work, which are presented as follows:

1. Although the workload model introduced in this thesis is comprehensive for the simulation use of parallel scheduling because it provides adequately three workload attributes (including arrival time, runtime and number of processors), the model should include another attribute, namely the user estimated runtime, since this attribute is essential for backfilling scheduling. As the user estimated runtime attribute has a strong correlation with the other attributes, modeling it separately would be impractical and unrealistic. Instead, it should be modeled in conjunction with the Bag-of-Tasks behaviour and temporal locality because it has been demonstrated that similar jobs often exhibit similar estimates [105].
2. The importance of periodicity such as daily cycles is demonstrated in [22]. Therefore, this feature should be captured and integrated into the workload model.

3. Job failures/cancels are also interesting for consideration in any workload model [52]. The occurrence of a job failure/cancel event can release computing resources, stimulate schedulers to re-schedule jobs, and so apparently affect scheduling. Despite their importance, a model for job failures/cancels is still lacking in the literature.
4. Research on performance evaluation of workload features is still scarce in the literature. In particular, the behaviour of workload features on grids is an open area.
5. Though starting a long time ago, parallel system scheduling is a hard problem that is still actively researched now [26, 27]. Compared with cluster scheduling, grid scheduling is a broader subject [77]. Despite the large number of scheduling strategies for both clusters and grids, these strategies are often designed without taking into account performance impacts of workload features, and thereby miss important clues that can improve schedulers. Hence, scheduling design based on workload features is a promising research direction for future work.