

CHAPTER 6

CONCLUSION AND FUTURE WORK

This thesis has presented a novel strategy to schedule workflow-based parameter-sweep applications onto clusters. It is aimed to make the workflow that gives the best result finishes as soon as possible. The technique is based on evaluation of intermediate results and dynamic priority adjustment. The preliminary evaluation of the technique in comparison with a conventional scheduling strategy on a drug screening application has shown some promising results. The scheduling was implemented on top of a batch job scheduler called Torque. However, many assumptions are made for simplicity and therefore need more experiments in order to evaluate the applicability of the algorithm in general.

Therefore, it was tested on a larger simulation environment with generated workload. The experiment also explore the effects of some parameters, such as the choice of P_1 and P_2 , and the effectiveness of the algorithm for ranges of application characteristics, namely the correlation between the quality of intermediate results and of final results, and the fraction of the computation that is required to obtain final results from intermediate results. The results indicate that the proposed algorithm meets the objectives. It also works better than other algorithms, when the correlation between the intermediate result and final result is more than 0.5 and when the evaluation point is in the range of 30% to 80%.

In the future, we plan to investigate the technique with on-line submission of a larger number of more complicated workflows on a larger cluster. Also, we are improving the scheduler to work on the Grid.