

# A Simulation Toolkit to Investigate the Effects of Grid Characteristics on Workflow Completion Time

**Title:** [A Simulation Toolkit to Investigate the Effects of Grid Characteristics on Workflow Completion Time](#)

**Authors:** M. O. McCracken, A. Snavely

**Abstract:** Advances in technology and the increasing number and scale of compute resources have enabled larger computational science experiments and given researchers many choices of where and how to store data and perform computation. Analyzing the time to completion of their experiments is important for scientists to make the best use of both human and computational resources, but it is difficult to do in a comprehensive fashion because it involves experiment, system and user variables and their interactions with each configuration of systems. We present a simulation toolkit for analysis of computational science experiments and estimation of their time to completion. Our approach uses a minimal description of the experiment's workflow, and separate information about the systems being evaluated.

We evaluate our approach using synthetic experiments that reflect actual workflow patterns, executed on systems from the NSF TeraGrid. Our evaluation focuses on ranking the available systems in order of expected experiment completion time. We show that with sufficient system information, the model can help investigate alternative systems and evaluate workflow bottlenecks. We also discuss the challenges posed by volatile queue wait time behavior, and suggest some methods to improve the accuracy of simulation for near-term workflow executions. We evaluate the impact of advance notice of predictable spikes in queue wait time due to down-time and reservations. We show that given advance notice, the probability of a correct ranking for a sample of synthetic workflows could increase from 59% to 74% or even 79%.

**Reference:** @conference{mccracken2009simulation, title={{A simulation toolkit to investigate the effects of grid characteristics on workflow completion time}}, author={McCracken, M.O. and Snavely, A.}, booktitle={Proceedings of the 4th Workshop on Workflows in

Support of Large-Scale Science}, pages={1--10}, year={2009},  
organization={ACM} }