Panorama: Modeling The Performance of Scientific Workflows

Ewa Deelman[1], Christopher Carothers[4], Anirban Mandal[3], Brian Tierney[2], Jeffrey Vetter[5], Ilya Baldin[3], Claris Castillo[3], Gideon Juve[1], Dariusz Król[1,6], Vickie Lynch[5], Ben Mayer[5], Jeremy Meredith[5], Thomas Proffen[5], Paul Ruth[3], Rafael Ferreira da Silva[1]

[1]Information Sciences Institute, University of Southern California
[2]Lawrence Berkeley National Laboratory
[3]RENCI/UNC Chapel Hill
[4]Rensselaer Polytechnic Institute
[5]Oak Ridge National Laboratory
[6] AGH University of Science and Technology, Department of Computer Science, Krakow, Poland

Abstract

The Panorama project aims to further the understanding of the behavior of scientific workflows as they are executing in heterogeneous environments. Panaroma's approach to modeling and diagnosing the runtime performance of complex scientific workflows is to integrate extreme-scale systems testbed experimentation, structured analytical modeling and parallel systems simulation into a comprehensive workflow framework that can characterize the end-to-end workflow performance on today's and future generation architectures, which can be used to improve the overall workflow performance and reliability. We first present the Panaroma architecture, including the individual framework components: the Aspen analytical application modeling software, the ROSS simulation framework, the Pegasus workflow management system and how they are used to model the behavior of DOE-relevant applications. We then describe how analytical models can be augmented with detailed simulations. By having a coupled model of the application and execution environment, decisions can be made about resource provisioning, application task scheduling, data management within the application, etc. Finally, our approach for correlating the real time application and infrastructure monitoring data is presented and how it can be used to verify application behavior, perform anomaly detection and diagnosis, and support adaptivity during workflow execution.

Acknowledgments

This work was funded by DOE under contract #DE-SC0012636, "Panorama - Predictive Modeling and Diagnostic Monitoring of Extreme Science Workflows". The development of the neutron scattering simulation workflow was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division. The use of Oak Ridge National Laboratory's Spallation Neutron was sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences.