lecture image

CCT Colloquium Series

Challenges of Managing Large-Scale Scientific Workflows in Distributed Environments

Ewa Deelman, USC Information Sciences Institute
Research Assistant Professor
Johnston Hall 338
September 29, 2006 - 03:00 pm

Abstract:
Recently, workflows have emerged as a paradigm for conducting large-scale scientific analyses. The structure of a workflow specifies what analysis routines need to be executed, the data flow amongst them, and relevant execution details. These workflows often need to be executed in distributed environments, where data sources may be available in different physical locations and the processing steps may have different execution requirements. Workflows help manage the coordinated execution of related tasks. They also provide a systematic way to capture scientific methodology and provide provenance information for their results. Scientists in many disciplines are approaching data volumes and resource sharing facilities that would enable a new stage in scientific discovery. Although many advances have been made to enable scientists to efficiently and easily use workflow technologies, many challenges remain. Based on our experiences in generating and managing workflows with thousands of tasks that execute over 1.8 CPU/years and process approximately 10 TB of data, this talk will describe open research problems in workflow management. We will explore the challenges from the perspective of workflow creation, compilation and execution and describe advances in each area.

Speaker's Bio:
Ewa Deelman is an Research Assistant Professor at the USC Computer Science Department and a Research Team Leader at the Center for Grid Technologies at the USC Information Sciences Institute. Dr. Deelman's research interests include the design and exploration of collaborative scientific environments based on Grid technologies, with particular emphasis on workflow management as well as the management of large amounts of data and metadata. At ISI, Dr. Deelman is leading the Pegasus project, which designs and implements workflow mapping techniques for large-scale workflows running in distributed environments. Pegasus is being used day-to-day by scientists in a variety of disciplines including astronomy, gravitational-wave physics, earthquake science and many others. Dr. Deelman received her PhD from Rensselaer Polytechnic Institute in Computer Science in 1997 in the area of parallel discrete event simulation. Dr. Deelman is an Associate Editor responsible for Grid Computing for the Scientific Programming Journal and a chair of the GGF Workflow Management Research Group.

Center for Computation & Technology
2003 Digital Media Center • Telephone: +1 225/578-5890 • Fax: +1 225/578-8957
© 2001–2025 Center for Computation & Technology • Official Web Page of Louisiana State University.