Chapter 4

Collaborative and Distributed e-Research Environment for Supporting Scientific Research and the Education Process

Dukyun Nam
Korea Institute of Science and Technology Information, Republic of Korea

Junehawk Lee[1]
Korea Institute of Science and Technology Information, Republic of Korea

Kum Won Cho
Korea Institute of Science and Technology Information, Republic of Korea

ABSTRACT

The efficient use of a scientific application service built on a computing environment requires technology that integrates each application service into a workflow so that the workflow is executed in a cooperative environment. There have been a number of attempts to automate research activities as a scientific workflow. However, there are practical problems in the full automation of research activities for a number of simulation programs and researchers. In the cyber environment for Collaborative and Distributed E-Research (CDER), the types of workflows need to be studied and supported separately and with different methodologies. In this chapter, the authors analyze the scientific research and education processes and categorize them into four types: simulation, experiment, collaborative work, and educational activity. They then describe the applications needed for each category. To justify their categorization of the CDER workflow, they examine the workflow of e-AIRS (e-Science Aerospace Integrated Research System), a problem-solving environment for aerospace research.

DOI: 10.4018/978-1-4666-0125-3.ch004

INTRODUCTION

In terms of computational science, the Collaborative and Distributed E-Research (CDER) environment is similar to the science gateway in e-Science defined as "large scale science that will increasingly be carried out through distributed global collaborations enabled by the Internet." Science Gateway (Wilkins-Diehr, 2007) provides not only an access interface for computing resources, information, and instruments, but also an intelligent environment, i.e., a cyber environment (Myers & McGrath, 2007), wherein users can retrieve useful information and use research tools. Through this interface, users are able to conduct computer simulations to verify and understand existing or proposed theories. The efficient use of a scientific application service built on a computing environment requires technology that integrates each application service into a workflow so that the workflow is executed in a cooperative environment. The business community has tried to automate business processing steps, and the computer industry supplied tools to help with this. In e-Science, research activities have been automated as a scientific workflow (Barga & Gannon, 2007; Leymann & Roller, 2000). However, the full automation of research activities with a number of stand-alone software programs has practical problems. For the cyber environment for Collaborative and Distributed E-Research (CDER), we need to study the types of workflows and support each type separately with different methodologies.

In this chapter, we analyze the research flows for CDER and describe the essential functionality of the workflow system, even when the system does not fully automate the entire research process. To do this, the workflow expected in CDER is categorized into four types. The first type describes the workflow for simulations on remote computing resources that involve supercomputers or grid resources. The second type is for experiments involving large-scale equipment. The third type describes the workflow for collaborative work. Finally, the fourth type describes workflow for educational activity. The CDER can assist students in understanding natural phenomena in class in areas such as computational fluid dynamics, bioinformatics, and computational chemistry. The educational activities (Kim et al., 2006) have different characteristics to research activities. For example, in an educational activity, students might run a sample simulation simultaneously for a limited time to acquire knowledge, including using a science gateway.

To justify our categorization of the CDER workflow, we examine the workflow of e-AIRS (e-Science Aerospace Integrated Research System) (Kim, et al., 2006), a problem-solving environment for aerospace research. We explain the characteristics of the CDER workflow, which benefits business and scientific workflows, based on the analysis of the e-AIRS system. Here, we also describe the necessary applications and functionalities for the workflow.

RELATED WORK

Directed Acyclic Graph Manager (DAGMan) (Frey, 2002) allows users to submit a number of jobs with workflows using Condor, and provide the interface for various types of workflows. The goal of DAGMan is the automation of managing complex workflows, including the job submission process. The weakness of DAGMan is the lack of supporting control flow, such as conditional branch and iteration, because Directed Acyclic Graph (DAG) is limited in how it represents the dependency of each step. Pegasus (Deelman, et al., 2003), using the DAGMan execution engine, constructs executable workflows based on the information of workflow instances and usable resources. It makes it possible for users to design a workflow at the application level, regardless of the status of computing resources and the execution environment.

Related Content