DOE Science Grid: Enabling and Deploying the SciDAC Collaboratory Software Environment

Principal Investigators: William E. Johnston, Lawrence Berkeley National Laboratory, Ray Bair, Pacific Northwest National Laboratory, Ian Foster, Argonne National Laboratory, Al Geist*, Oak Ridge National Laboratory, William Kramer, National Energy Research Scientific Computing Center. Science Grid Working Group: Keith Jackson, LBNL, Tony Genovese, ESnet, Mike Helm, ESnet, Von Welch, ANL, Steve Chan, NERSC, Kasidit Chanchio, ORNL, Scott Studham, PNNL

Summary

DOE's large-scale science projects involve many collaborators at multiple institutions. The DOE Science Grid is being developed and deployed across the DOE complex to provide persistent Grid services to advanced scientific applications and problem solving frameworks. Fault monitoring and auditing are two key features necessary for the production-quality Grid computing. Mechanisms and software tools providing these capabilities have been developed and evaluated. Initial results are promising and showing usefulness of the tools.

The DOE Science Grid project is a collaboration of five DOE laboratories under the SciDAC program to create an advanced cyber infrastructure that enable geographically separated scientists to effectively work together as a team and to facilitate remote accesses to both facilities and data. The goal of the project is to provide persistent and scalable Grid services to support DOE large-scale science projects. These services provide security, resource access and management functionalities to advanced scientific applications and problem solving frameworks.

Fault monitoring and auditing are key Grid services required for the construction of production-quality DOE Science Grid cyber infrastructure. Use of high-performance applications across the DOE Science Grid will require access to accurate, up-to-date information on the structure and state of available resources. For example, are all the resources required by the application available at this moment? Are the resources free for the entire period of the run or are there conflicts downstream? The DOE Science Grid information services and scheduling tools will be available to SciDAC researchers for determining the best configuration based on the information available at the start of a job. *However, the present Grid infrastructure is missing the monitors and tools to dynamically assess its resources and respond*. The resources across the DOE Science Grid are constantly changing—nodes fail in a compute resource, a drive fails in a storage archive, the power outages, etc. Wide spread acceptance of the DOE Science Grid as a viable computational resource requires the creation and deployment of appropriate monitoring and auditing tools.

Our accomplishments to date are:

Grid monitoring: ORNL has developed a prototype fault monitoring system to

* Phone: 865-574-3153 gst@ornl.gov

monitor availability of computing resources as well as Grid services running on them. A novel and practical leader-election protocol has been developed and incorporated into the fault monitoring system to provide consistent and reliable monitoring capability over distributed resources despite computer or network failures. The system also features job monitoring functionality and can interoperate with Grid Information Services (GIS). The system has been demonstrated at Super Computing 2002 conference and recently applied to monitor availabilities of mass storage resources and Grid services for the Earth System Grid (ESG) project.

Grid resource accounting: An investigation was done on Grid Resource Usage Data and Accounting (RUDA) requirements and on current existing accounting system examples of various computers. Based on the investigation, a model of RUDA was designed. A prototype has been developed, and the first version is undergoing test runs on the ORNL Grid test bed. This model is a distributed system where each resource entity runs a lightweight daemon server, which can be configured into two modes. One mode manages resource usage data and accounting of an individual resource entity. The other mode serves clusters and can also be used to monitor and manage resource usage data and accounting of any interested resource group on a Grid. The RUDA information can be conveniently accessed through web or client APIs. The design emphasizes flexibility so that it can be integrated into other Grid software or customized to fit requirements of different sites.

Both fault monitoring and RUDA systems are designed to be efficient, reliable and easy to use. They have made impacts in fault tolerance and resource management areas.

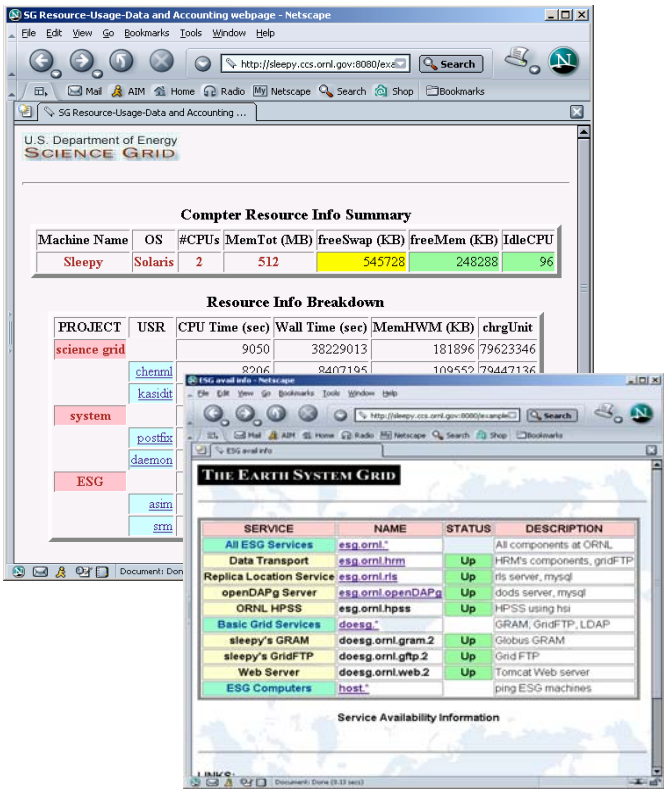

Figure 1. Example screenshots showing resource utilization and accounting information reported by RUDA system (top), and availability status of the ESG services reported by the fault monitoring system (bottom).

Their contributions are initial steps toward the construction of persistent, reliable production Grid computing environments. We are currently developing fault notification mechanisms on top of the fault monitoring system. In the next step, we will work with other SciDAC scientists to investigate their fault monitoring and accounting requirements and customized our tools to fit their needs.

For further information on this subject contact:
Mary Ann Scott, Program Manager
Mathematical, Information, and Computational Sciences Division
Office of Advanced Scientific Computing Research
Phone: 301-903-6368
scott@er.doe.gov