Commodity Grid Kits – *Making Grids Easy to Program and Use*

Gregor von Laszewski, ANL
Keith Jackson, LBL

Summary

Many state-of-the-art scientific applications need numerous distributed resources to fulfill the quest of advancing multidisciplinary scientific research. The Grid provides an infrastructure that can be used as part of this scientific quest. Our project promotes the use of commodity technologies, already in use by thousands of scientists, to access the Grid from higher-level programming frameworks. These frameworks form the basis for scientific portals, promoting collaboration between large scientific teams, and allow the harnessing of distributed resources that are needed for solving challenging scientific problems.

The computing demands of many state-of-the-art scientific applications—such as climate modeling, astrophysics, high energy physics, structural biology, chemistry, and tele-immersive engineering—requires the coordinated use of numerous distributed and heterogeneous components, including advanced networks, computers, storage devices, display devices, and scientific instruments. Such a national collaborative Grid infrastructure is being developed and supported by DOE, NSF, and NASA.

Developing advanced scientific *applications* for these emerging national-scale computational Grids is still a difficult task, however. Although elementary Grid services exist that enable scientific application developers to authenticate, access, manage, and discover sophisticated remote resources, these frameworks are not compatible with the commodity technologies and frameworks used by application scientists today. Additionally the technologies often are too complex to be used by computational scientists, who may lack the expertise needed for dealing with such a complex infrastructure. A higher level of abstraction is demanded that supports easy access to the Grid within frameworks that are, and will be, used for scientific problem solving.

We are developing Commodity Grid (CoG) Kits that allow the scientific application programmer or middleware developer to readily make use of Grid services from a higher-level framework. We focus this effort on the development of two CoG Kits: one for Java and the other for Python. These kits will allow for easier and more rapid application development. Based on component models, they encourage collaborative code reuse among problem-solving environments, science portals, Grid middleware, and collaboratory pilots currently under development.

Our work is currently used in more than a dozen scientific research projects, including projects sponsored by NSF, NIH, NASA, and DOE. Moreover, we collaborate closely with the scientific community and other DOE-sponsored activities to define scientific portals based on the technology we provide. Indeed, CoG Kits have become the de facto community standard for developing portal applications accessing the Grid.

As we remove a large portion of the burden of dealing with the complex Grid infrastructure, research teams can concentrate on the scientific aspects of their problems. Examples of successful applications of our technologies can be

found in the development of portals for climate and atmospheric modeling, for astrophysical black-hole simulations, and for structural biology. Furthermore, our technology is an integral part of the development of the DOE Science Grid, allowing the larger DOE science community to utilize an emerging DOE-wide resource infrastructure accessible through Grid technologies.

We highlight here just two of the many examples of the use of our code. Our technology is used as part of the Access Grid (Figure 1), an ensemble of resources that can be used to support human interaction across the Grid. Through the Access Grid, scientists can collaborate easily in virtual meeting rooms.

Figure 1. The Access Grid uses the Python CoG Kit (pyGlobus) as a major component to enable scientific collaborative sessions.

Our technology is also used in an NSF-sponsored atmospheric simulation tool. In this project a Grid-based portal is used to analyze the study of the April 1996 storm that struck Illinois and nearby states with high winds and strong tornadoes. Ensemble modeling studies of this case involve hundreds of simulations to span the range of observed convective environments and thunderstorm cell behavior. Our technology enables the scientific application developer to use the Grid for this purpose.

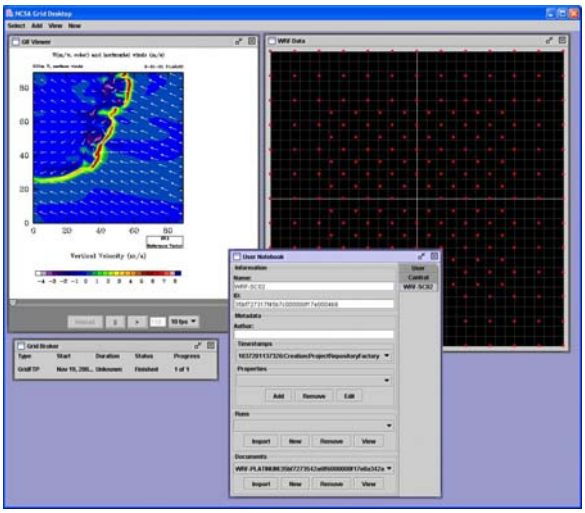

Figure 2. Our technology is also being used by researchers outside of DOE for building sophisticated portals to Grids for atmospheric simulations (NCSA WARF).

The next step is to incorporate advanced Grid services into the CoG Kits and to develop a high-level abstraction layer that can be readily integrated into workflows targeted for Grid applications. To achieve our ambitious goals, we will collaborate with information technology and application teams. We anticipate that many scientific projects will use our technology in order to build application specific portals to Grid-enabled applications.

For further information on this subject contact:

Gregor von Laszewski, Principal Investigator
Argonne National Laboratory
Mathematics and Computer Science Division
Phone: 630-252-0472
e-mail: gregor@mcs.anl.gov

Keith R. Jackson, Principal Investigator
Lawrence Berkeley National Laboratory
Distributed Systems Department
Phone : 510-486-4401
email : krjackson@lbl.gov