

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 as part of the scientific problem solving process.

Many state-of-the-art scientific applications, such as climate modeling, astrophysics, high energy physics, structural biology, chemistry, and tele-immersive engineering, require 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 Grid infrastructures is, however, still a difficult task. Although elementary Grid services exist that enable scientific application developers to authenticate, access, manage, and discover sophisticated remote resources, these services are not compatible with the commodity technologies and frameworks used by application scientists today. Additionally, many computational scientists lack the necessary expertise to deal with such a complex infrastructure.

To address this problem, we are developing Commodity Grid (CoG) Kits that will enable scientific application programmers and middleware developers 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. Based on component models, these kits encourage collaborative code reuse and avoid duplication of effort among problem solving environments, science portals, Grid middleware, and collaboratory pilot developers.

The CoG Kits provide compatibility to the Globus Toolkit through which terra Grid facilities can share resources among distributed domains. In Java we developed protocol compatible components interfacing with the Grid infrastructure, enabling us to develop sophisticated Web enabled clients and services. In Python we developed components that are making use of precompiled Globus Toolkit libraries. Both promote Grid programming within their frameworks making it today possible to develop interfaces and programs to the Grid in Java and Python.

The new CoG Kit technology has been reused while building a number of scientific portals to the Grid. Indeed, it already has become the de facto community standard for developing Grid portal applications. One exciting example is the NSF portal for atmospheric simulations, which is being used to analyze the April 1996 tornado that struck Illinois and nearby states (see Figure 1).

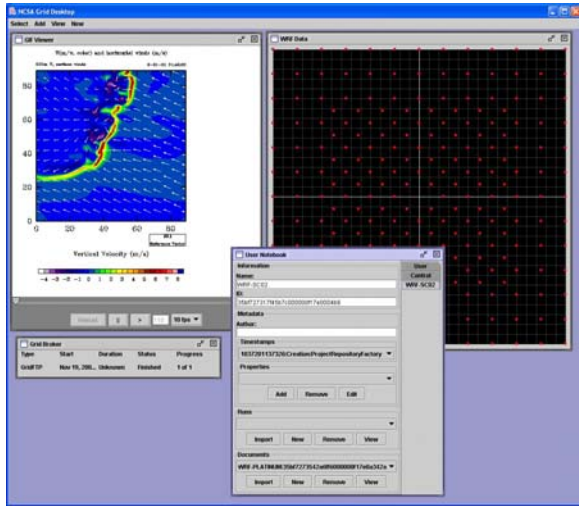


Figure 1. Our technology is already being used by researchers outside of DOE, for example in this sophisticated Grid portal for atmospheric simulations (NCSA WARF)

Modeling of this tornado involves hundreds of simulations of observed convective environments and thunderstorm cell behavior. Through this portal scientist are able to steer the simulation, stage the calculations and obtain data for the calculations from distributed resources located on the Grid.

Other examples of successful applications of our technologies include portals for astrophysical black-hole simulations, portals for structural biology, the SciDAC Earth System Grid, and the SciDAC Particle Physics Data Grid. Furthermore, our technology is integral part in the development of the current generation of a DOE-wide Science Grid allowing the larger DOE science community to utilize an emerging DOE wide Grid infrastructure.

CoG Kits have also been incorporated in the Access Grid (Figure 2), an ensemble of resources that enables group-to-group collaborations amongst interdisciplinary scientists. The CoG Kit provides an efficient means for authentication and secure communication between different rooms.



Figure 2. The Access Grid uses the Python CoG Kit (pyGlobus) as a major component to enable scientific collaborative sessions.

Furthermore, CoG Kits provide an elementary building blocks for developing the next generation Grid infrastructure as part of upcoming Globus Toolkit releases.

The SciDAC project has enabled us to work tightly as an interdisciplinary team between scientists and technology providers amongst a variety of SciDAC projects. Interdisciplinary interactions, access to supercomputing facilities, and new collaborative use of sophisticated resources are among the results of our group efforts.

In the next years, additional steps are needed that will allow us to use our basic middleware as part of high-end scientific collaborative environments while reusing methodologies and frameworks that make the construction of increasingly sophisticated collaborative environments possible. We will continue to collaborate with scientists to create transparent use of terrascale facilities by the scientist.

For further information on this subject contact:

Dr. 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 : 510-406-1407 email : krjackson@lbl.gov