Grid Tools for Applications

Gabrielle Allen* Thomas Dramlitsch* Ian Kelley* Gerd Lanfermann* Thomas Radke*
Michael Russell† Ed Seidel*‡ John Shalf§ Oliver Wehrens*

Cactus[1] is an open source problem solving environment designed for scientists and engineers. Its modular structure facilitates parallel computation across different architectures and collaborative code development between different groups. The Cactus Code originated in the academic research community, where it has been developed and used over many years by a large international collaboration of physicists and computational scientists. Cactus has a large userbase in the numerical relativity community, where physicists have developed many modules for evolving astrophysical phenonema such as black hole and neutron star collisions. The demanding and varied computational requirements for this problem has driven and guided the development of a number of Grid-enabled tools both to enable larger and higher fidelity simulations and to improve the working environment and practises of the community. In this demonstration we will show a selection of our current technologies [2] providing the following capabilities:

Accessibility to Resources: The *Cactus User Portal*, developed through the NSF ASC project, and now additionally the EU GridLab project, provides a easy-to-use interface for using remote machines, in both a local and Grid environment, providing code assembly, resource finding and authentication, executable staging, and simulation monitoring and steering. The Portal also acts as a collaborative tool, facilitating code sharing, tracking simulation status, and providing access to simulation output.

Remote File Access: Large-scale computer simulations generate large-scale data sets. Conventional analysis and visualization then becomes prohibitively resource-intensive when remote simulation data must be moved to a local machine for processing. Enhancements to the *Hierarchical Data Format* `HDF5` I/O library, allow existing I/O layers to operate directly on remote files which are uniquely addressed by their URL.

Remote Visualization, Monitoring and Steering: Remote visualization is the capability to visualize data (possibly in a virtual file streamed live from a running simulation) from a remote resource with a client on a local machine, eliminating the need to move enormous amounts of data between machines. Cactus provides several different implementations of data streaming. which can be viewed using various visualization tools (e.g. [3]). Scientists using large scale remote resources have to cope with different accounts, networks, operating systems, and queuing systems. Cactus alleviates many of these problems with a module which provides a simulation with its own webserver providing detailed simulation information viewable from any remote web browser. The web interface can also be used for steering: e.g. pausing or terminating a simulation, as well as changing the values of parameters, for example changing data output properties.

Dynamic and Distributed Grid Computing: Distributing a single simulation across multiple resources provides the means to run larger simulations than otherwise possible on a single machine, or to couple available machines for running simulations immediately, Cactus applications are easily distributed using the Globus Toolkit. We are developing a *Grid Application Toolkit* as part of the EU GridLab project to exploit the Grid for new dynamic application scenarios. Existing tools allow for automatic simulation migration and spawning of independent tasks to different (faster/cheaper/available) resources.

References

[1] Cactus Code homepage: `http://www.cactuscode.org`

[2] G. Allen, W. Benger, T. Goodale, H. Hege, G. Lanfermann, A. Merzky, T. Radke, E. Seidel, J. Shalf, "Cactus Tools for Grid Applications", *Cluster Computing*, **4**, 179-188, (2001), [`http://www.cactuscode.org/CacPapers/CactusTools.ps.gz`]

[3] Amira: `http://amira.zib.de`, IBM Data Explorer: `http://www.research.ibm.com/dx`, LCA Vision: `http://zeus.ncsa.uiuc.edu/ miksa/LCAVision.html`.

*Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, Golm
†University of Chicago, Chicago, IL
‡National Center for Supercomputing Applications, Champaign, IL
§Lawrence Berkeley National Laboratory, Berkeley, CA, (LBNL)