



IST-2001-32133

GridLab - A Grid Application Toolkit and Testbed

D2.3 Grid Enabled Cactus Applications

Author(s):	Gabrielle Allen
Document Filename:	Gridlab-2-D2.3-0003.CactusGAT
Work package:	WP2 - CactusGAT
Partner(s):	PSNC,MPG,ZIB,MU,VU,CARDIFF
Lead Partner:	MPG
Config ID:	Gridlab-2-D2.3-DRAFT-1
Document classification:	PUBLIC

Abstract: This document describes the current state of Cactus application grid enabled through the CGAT and GAT.



1 Overview

The goal of GridLab WP2 (CGAT) is to enable applications written in the Cactus framework to use of the Grid Application Toolkit (GAT) to support current production applications and new, sophisticated and advanced Grid related scenarios.

While the GAT-API provides all applications with abstract Grid capabilities, within a framework such as Cactus, these API's can be customized for a particular application, and used to build more advanced composite Grid capabilities — this is the purpose of the CGAT set of Cactus thorns described in Deliverable 2.4.

In this document we briefly list Cactus applications and scenarios which are making use of the GAT, either through the CGAT set of thorns, or through associated tools using the GAT-API directly. Note that this work is ongoing, and the GAT is being used to develop Grid scenarios for a wide range of Cactus applications.

2 Grid-enabled Cactus Scenarios

2.1 Data Management of Cactus Output Files

One of the most pressing needs for computational scientists, such as the numerical relativists at the AEI, is the ability to track and share the many output files from simulations. We are currently developing a Cactus thorn for the CGAT which will dynamically and automatically register output files, and associated metadata, with a replica catalogue. These files, and the metadata associated with them, can then be searched and accessed via a variety of mechanisms, including standalone client tools, the Cactus Portal, and the VizLauncher application described in this section.

2.2 Remote Visualization Tool for Cactus Output Data

The remote VizLauncher (shown in Figure 1) is a GAT-based utility that adds remote capabilities to a pre-existent visualization tool (<http://www.cactuscode.org/VizTools/VizLauncher.html>). The VizLauncher uses the path name of a file to be visualized and a mime-type description as input parameters to make automatic or user-assisted decisions about the visualization tool to be started locally in order to visualize the file. The remote VizLauncher adds the possibility to browse a metadata catalog where logical filenames for data generated by previously run simulations are stored together with attributes like simulation ID, the mime type of the file and a short textual description of the file. The tool takes a simulation ID as an input, it queries the metadata catalog for logical files that match the simulation ID using the GATAdvertService and launches a GUI browser that helps the user browse and select one logical file to be visualized. After the logical file name is determined, the GATLogicalFile is used to retrieve a physical name for the logical file which is then used by a GATFile adapter that downloads the file locally. The final step is to run the VizLauncher with the local file path and the MIME type retrieved from the GATAdvertService.

2.3 Cactus Task Farming Infrastructure

The Cactus Task Farm Infrastructure is a tool that enables loosely coupled, parallel distributed computations across a collection of computational resources. CTFI provides a simple API to support a broad range of task farming applications through predefined mechanisms to specify application specific file staging, input requirements, and individual task requirements. A set of pluggable components allow CTFI to leverage multiple technologies to provide basic functionality

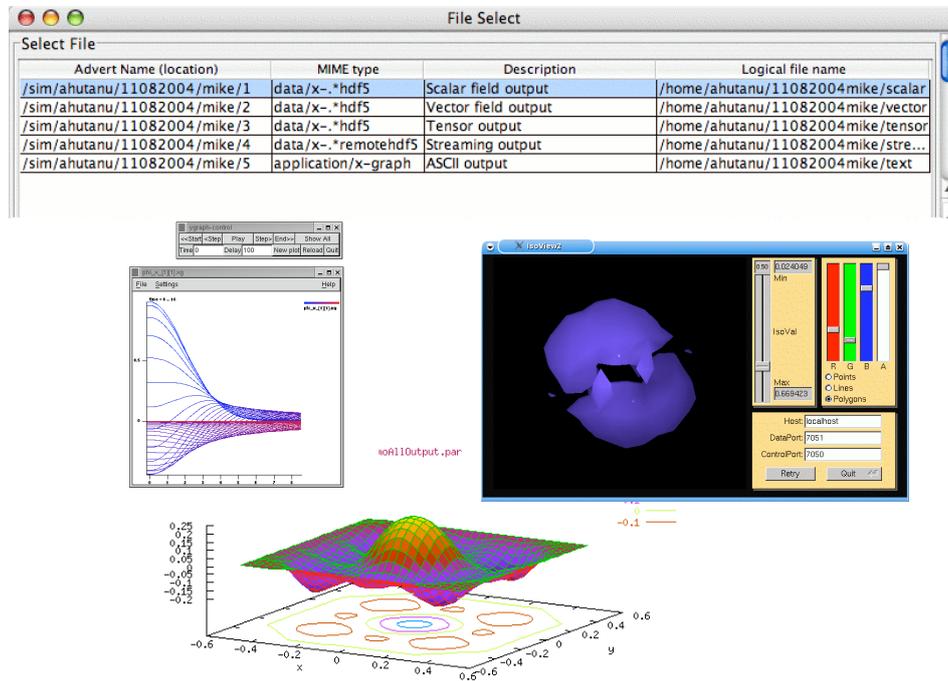


Figure 1: GAT-enabled VizLauncher

such as scheduling, file management, and resource discovery. CTFI is application independent and does not require recompilation of existing task applications to run. The CTFI is an ongoing effort by the Grid Computing group at the AEI and CCT.

The CTFI uses the CGAT to provide resource discovery (e.g. GRMS), file management (e.g. scp, gisftp, gridftp), and staging of the task farm managers (e.g. ssh, GRMS, GRAM). The CTFI is now being used for several external projects.

2.4 Migration of Cactus Applications

Migration of multiprocessor Cactus simulations between production clusters and supercomputers has been a driving scenario for the GridLab project. Motivated by pre-GridLab early prototypes such as the *Cactus Worm*, now in GridLab the emphasis has been on providing dependable and fault tolerant infrastructure to enable production use of migration for real world applications. Migration involves most other components developed in the GridLab project (Figure 2), and has helped coordinate the development of services.

From the experiences with Cactus Migration, we have also developed a wrapper to handle migration of legacy (non-Cactus) applications.

2.5 Cactus Portlets for Grid Sphere

The Cactus Portal is a web-based application built using the GridSphere framework. The portal has taken the general functionality of GridSphere and expanded it to include the application specific components necessary for Cactus. These components consist primarily of application

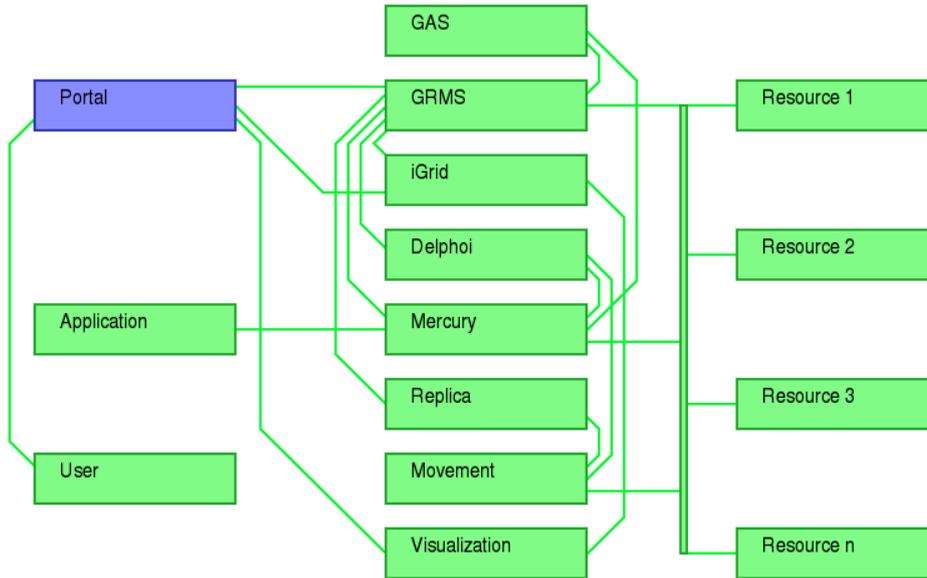


Figure 2: GridLab services employed for Cactus migration scenario

building, tracking, data management and job submission, as well as enhanced notification services that provide a user with advanced methods of monitoring their simulation's progress. The Cactus Portal uses multiple services from GridLab, including data management (logical file browser and remote data access), resource information, job submission (GRMS and GRAM). More portlets are currently being developed to add and browse meta data associated with logical files.