



IST-2001-32133

GridLab – A Grid Application Toolkit and Testbed

D13.5 “Report on the First Annual Workshop”

| | |
|-----------------|---|
| Author(s): | Jarek Nabrzyski |
| Title: | Report on the First Annual Workshop |
| Subtitle: | Year 2002 |
| Work Package: | 13 |
| Lead Partner: | PSNC |
| Partners: | ALL |
| Filename: | GridLab-13-D.5-0001-Report_AnnualWorkshop |
| Version: | 1.0 |
| Config ID: | GridLab-13-D.5-0001-1.0 |
| Classification: | IST |

Abstract: This document is a GridLab report on the first annual workshop.

Project Manager: Jaroslaw Nabrzyski
Institute of Bioorganic Chemistry PAS
Poznan Supercomputing and Networking Center
ul. Noskowskiego 12/14
61-704 Poznan, Poland
Phone:+48 61 858 2072, Fax: +48 61 852 5954
Email: naber@man.poznan.pl





Table of Contents

| | |
|---|----|
| 1. Introduction..... | 3 |
| 2. List of Participants and the Meeting Agenda..... | 3 |
| 3. Technical Achievements of the GridLab Annual Workshop..... | 8 |
| 4. Summary..... | 12 |

1. Introduction

The first annual GridLab Workshop was organized by PSNC. The event took place in Zakopane, on 12-22 September, 2002. The workshop was divided into two parts: GridLab Conference (open for public) and GridLab Workshop (internal meeting). The workshop was attended by 41 participants from GridLab. As it was the first workshop of a kind we decided to invite only some external people to take part in the Conference. These external people included some persons from the Crossgrid project, from the Clusterix project (polish national project) and some people were representing NCSA (USA). These people were invited as the first potential collaborators of the GridLab project. In general, the very first two days were organized in several sessions with GridLab presentations to give an update per every workpackage and every institution. The remaining time was devoted to the inter-workpackage meetings, API and interface negotiations and discussion and some other technical meetings.

The aim was to deliver the first prototypes. Note that the first prototypes have been scheduled for the end of month 12 but in general we wanted to be ready by Supercomputing 2002.

2. List of Participants and the Meeting Agenda

The following is the list of the registered participants:

| Person | Affiliation |
|---------------------|---|
| Thorsten Schuett | ZIB |
| Dave Angulo | University of Chicago |
| Kurowski, Krzysztof | PSNC |
| Kelly Davis | MPG |
| Ian Kelley | AEI |
| Felix Hupfeld | ZIB |
| Thilo Kielmann | Vrije Universiteit, Dept. of Computer Science |
| Andrei Hutanu | ZIB |
| Tomasz Piontek | PSNC |
| Marcin Adamski | PSNC |
| Piotr Grabowski | PSNC |
| Radoslaw Strugalski | PSNC |
| Jaroslav Nabrzyski | PSNC |
| Bogdan Ludwiczak | PSNC |
| Juliusz Pukacki | PSNC |
| Bartek Lewandowski | PSNC |
| Tomasz Ostwald | PSNC |
| Rob van Nieuwpoort | Vrije Universiteit Amsterdam |
| Zbyszek Krzewinski | PSNC |
| Jason Maassen | Vrije Universiteit Amsterdam |
| Ludek Matyska | Institute of Computer Science MU |
| Massimo Cafaro | University of Lecce |
| Italo Epicoco | University of Lecce |
| Serena Pati | University of Lecce |

| | |
|---------------------|--|
| Silvia Mocavero | University of Lecce |
| Maria Mirto | University of Lecce |
| Miroslav Ruda | MU |
| Kashif Rasul | Albert-Einstein-Institute |
| Michael Russell | Max-Planck-Institute-Albert-Einstein-Institute |
| Zoltán Balaton | MTA SZTAKI |
| Gábor Gombás | MTA SZTAKI |
| Petr Holub | Institute of Computer Science, Masaryk University Brno |
| Ian Taylor | Cardiff University |
| Martin Kuba | Institute of Computer Science, Masaryk University Brno |
| Ian Wang | Cardiff University |
| Edward Seidel | yes |
| Gabrielle Allen | Max Planck Institute for Gravitational Physics |
| Oliver Wehrens | AEI |
| Jason Novotny | AEI |
| Ruxandra Bondarescu | National Center for Supercomputing Applications |
| Marian Bubak | ACK Cyfronet, POLAND (Gridstart Project) |
| Greg Daues | NCSA - University of Illinois at Urbana-Champaign |
| Pawel Kuczynski | Czestochowa University of Technology (Clusterix project) |
| Tomasz Kuczynski | Czestochowa University of Technology (Clusterix project) |

The meeting agenda was as follows:

12th of September - Arrivals from Krakow to Zakopane by a bus shuttle

13-14th of September Friday - GridLab Conference / Agenda

16-20th of September - GridLab Workshop

21st of September - Closing Summary Sessions

22nd of September - Departure from Zakopane to Krakow

Conference Agenda (13th - 14th September 2002)

Friday, 13th September 2002

| | |
|---------------|--|
| 8:30 - 9:00 | Registration |
| 9:00 - 9:15 | Welcome and Opening (Jarek Nabrzyski) |
| 9:15 - 9:45 | GridLab Project Summary (Jarek Nabrzyski/Ed Seidel) |
| 9:45 - 10:15 | CrossGrid Project - Invited Talk by Marian Bubak (ACK CYFRONET) |
| 10:15 - 10:45 | GridLab WP5 Testbed (Ludek Matyska/Mirek Ruda) |
| 10:45 - 11:15 | Coffee Break |
| 11:15 - 11:45 | GridLab WP1 GAT (Tom Goodale) |
| 11:45 - 12:15 | GridLab WP2 CGAT (Tom Goodale) |
| 12:15 - 12:45 | GridLab WP6 Security (Tomasz Ostwald) |
| 12:45 - 13:15 | GridLab WP10 Information Services (Massimo Cafaro) |
| 13:15 - 15:00 | Lunch |
| 15:00 - 15:30 | GridLab WP11 Monitoring (Zoltan Balaton) |
| 15:30 - 16:00 | GridLab WP12 Mobile users (Florian Schintke/Bartosz Lewandowski) |



| | |
|---------------|--|
| 16:00 - 16:30 | European Distributed Supercomputing Network (EDSN), Eol Submitted to FP6 (Ed Seidel) |
| 16:30 - 17:00 | Coffee Break |
| 17:00 - 18:30 | Steering Committee and Technical Board Meeting |
| 20:00 - | Regional Folk Dinner |

Saturday, 14th September 2002

| | | |
|---------------|---|-----------------------|
| 9.00 - 9.30 | GridLab WP3 TGAT (Ian Taylor) | |
| 9.30 - 10.00 | GridLab WP4 Portals (Michael Russe!) | |
| 10.00 - 10.30 | GridLab WP7 Adaptive environment (Thilo Kielmann) | [pdf] |
| 10.30 - 11.00 | GridLab WP8 Data (Andre Merzky) | [pdf] |
| 11:00 - 11:30 | Coffee Break | |
| 11:30 - 12:00 | GridLab WP9 Resource Management (Juliusz Pukacki) | [pdf] |
| 12:00 - 12:30 | GridLab WP13 Exploitation and Dissemination (Jarek Nabrzyski) | |
| 12:30 - 13:00 | GridLab WP14 Administrative and Financial Coordination (Zbyszek Krzewinski) | |
| 13:00 - 15:00 | Lunch | |
| 15.00 - 15.30 | SOAP/WSDL/OGSA Services and ClassAd Resource Selection (Dave Angulo) | [pdf] |
| 15.30 - 16.30 | Preparation of GridLab Workshop | |
| 16:30 - 17:00 | Coffee Break | |
| 17.00 - 18.30 | Preparation of GridLab Workshop | |

During GridLab Conference we presented current state of the GridLab project as well as discussed further steps to be taken. The conference also included: Steering Committee meeting, Administrative and other project issues.

As one can see above all the WPs presented their activities and status. We had also a talk given by our invited guest Marian Bubak from the Crossgrid project.

GridLab Workshop

The main goal of GridLab Workshop was to organize several meetings among workpackages in small working groups. Right after the GridLab Conference, it was a great opportunity to have some technical discussions.

The goal has been carefully planned to involve all the workpackages. We have decided to develop and implement the scenario based on the idea of the job migration because of bad performance and using the GAT (Grid Application Toolkit) with all the underlying GridLab services. This scenario has been defined in the following way:

```
Job starting then Migration because of bad performance - final version
```

```
-----  
User starts job via portal, and gets notified when the job starts.
```

```
User then monitors GAT application through the portal and discovers that  
it is performing badly. The user requests migration to a system where  
the application will perform better.
```



In the below security is implicitly involved in all communications.

(First stage - user asks portal to start job)

User <--> portal

portal <--> MDS (where is resource manager)

portal <--> resource manager

resource management <--> MDS (list of resources)

resource management <--> adaptive

(resource management provides a list of candidate sites to the adaptation component)

adaptive <--> monitoring system

(adaptation component queries the monitoring system at each candidate site about "execution time information".

This might be either: a "unified speed measure" a.k.a. BogoMIPS

- smallest possible precision

- + but not that bad in practice

or: the result of an application-specific

- micro benchmark

- + quite expressive for the actual case

- application needs to provide this code

or: some prediction over previous runs of the same application

- + might be the best thing to do

- this historical data needs to be stored

somewhere

- if there is no data, resort to BogoMIPS)

Adaptation component returns a ranked list of suitable candidate sites to resource management

resource management <--> replication (exe)

replication <--> file movement

resource management <--> replication (input)

replication <--> file movement

resource manager (submits job)

resource manager <--> monitoring (here's a job to monitor with gridlab id #1, local #2)

(returns gridlab id #1)

portal <--> resource manager (where is monitoring service for #1)



portal <--> monitoring service registers for notification event for #1)

monitoring service <--> portal (job #1 has started)

portal <--> notification

[application <--> monitoring (here is my process id #3, my gridlab id is #1)]

(Second stage - user asks portal for performance info)

User <--> portal

portal <--> resource manager (where is mon service for #1)

portal <--> monitoring system (how is #1 performing)

monitoring system <--> application

(At this stage the user has information about the job and decides the job must be migrated)

(Third stage - user asks portal to migrate application)

User <--> portal

Portal <--> resource management (migrate job #1)

resource management <--> MDS (list of resources, as above)

resource management <--> adaptive

adaptive <--> monitoring system

(At this stage the resource management system has found a new resource to run the job on)

resource management <--> application (checkpoint)

application (checkpoints)

application <--> replication (here are my checkpoint files)

application (stops)

resource management <--> replication (exe)

replication <--> file movement

resource management <--> replication (input)

replication <--> file movement

resource management <--> replication (checkpoint files)

replication <--> file movement

(All relevant data has now been moved to new hardware resource)

resource management (submits job)

This scenario, including the mobile user support (notification) has been implemented using the Web Services technology. All the services and GAT are the pilot versions of the GridLab project deliverables. Demo of these was presented at the SC2002 as well as at the 1st GridLab review and it was very successful. Let us describe in some details the meeting outcome.

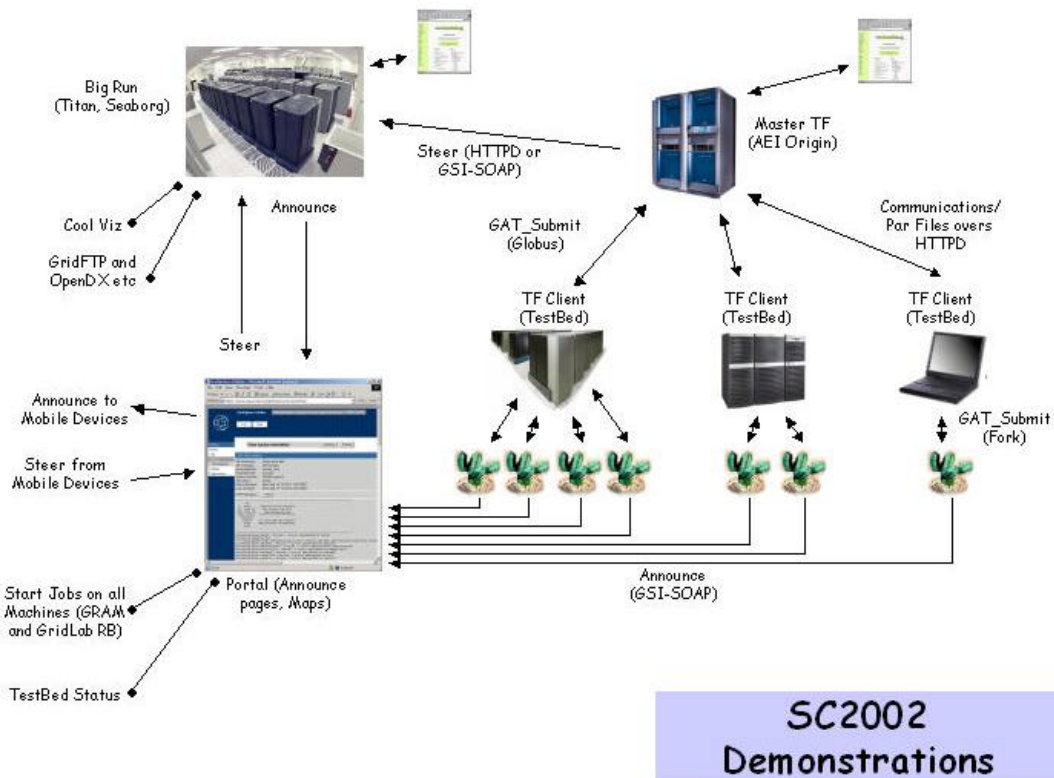
3. Technical Achievements of the GridLab Annual Workshop

This scenario was continuously being developed in order to be presented at the Supercomputing 2002 in Baltimore. By the end of the year the first prototypes of all the services and the GAT itself were ready. The future months will be devoted to full testing of the first prototypes and preparation of the next all-hands meeting in Eger, Hungary. During the meeting the lessons learned from the first prototype will be discussed, as well as the new, more ambitious scenario will be designed and implemented.

The GridLab team has been very successful with the SC2002 demonstrations. Our demonstrations showed how Grid techniques can be used to enhance both the effectiveness of large scale simulations, and the working environments of the computational physicists running them.

We used the black hole simulations of the [AEI numerical relativity group](#), using the [Cactus framework](#), and make use of the world wide [GGTC testbed](#) enabled by the [Globus toolkit](#). We showed how a large black hole simulation can be automatically steered by the results of a large number of exploratory tasks run across the testbed. The results of the large simulation were visualized by various remote tools, making use of gridftp functionality with OpenDX, the Vizapult software, and a range of other tools. The simulations are tracked, monitored and interacted with via the Cactus Portal, which also notifies users via email and SMS of significant events. The [GridLab](#) Grid Application Toolkit (GAT) provided the simulations with a generic API which accesses grid services, such as the GridLab Resource Management System and reliable file transfer services.

These demonstrations were entries for the [Bandwidth Challenge](#) and the [HPC Awards](#). We have won these challenges, although this was done within a collaboration of Global Grid Testbed Collaboration initiative.



The Global Grid Testbed Collaboration had about 70 machines on 5 continents and 14 countries of every conceivable architecture and OS type, from Sony Playstation 2 to 1024 processor linux clusters. Some statistics is available on [testbed page](#).

We demonstrated 4 different applications running on this collection of machines, some of which are described at <http://sc2002.aei.mpg.de>.

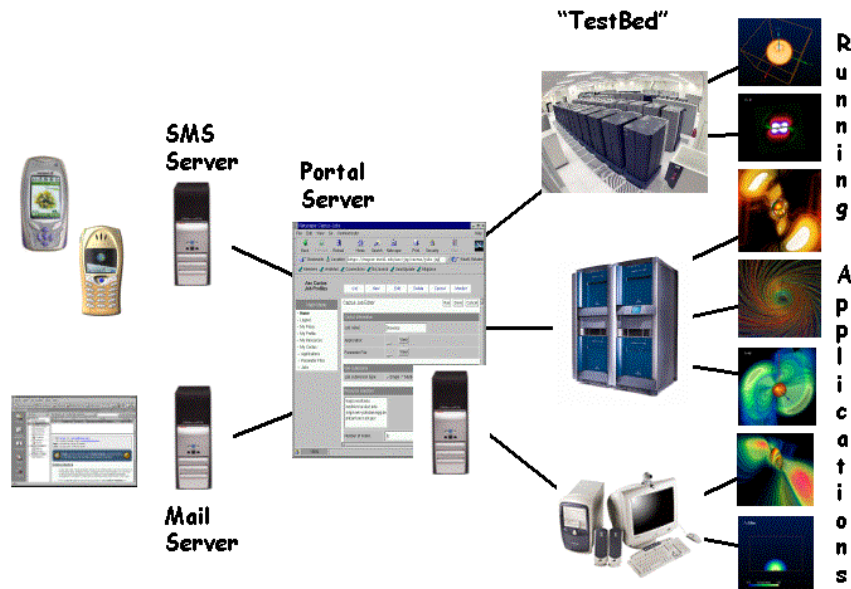
We entered both the HPC Challenge Competition, and in collaboration with John Shalf, Wes Bethel and others, we were also associated with an entry to the Bandwidth Challenge competition. We won 2 of 3 HPC Challenges (**the most geographically distributed** and **most heterogeneous applications**) AND John, Wes and team won **the Bandwidth Challenge competition by sustaining 16.8Gbits/sec** of data from a Cactus simulation running remotely (from various sites, from NCSA to Masaryk University), to the show floor where it was volume rendered on the fly!

Participating sites in the winning effort were the Parallel Distributed Systems Facility at the National Energy Research Scientific Computing (NERSC) Center at LBNL; and clusters at the SC2002 conference in Baltimore, Argonne National Laboratory, the National Center for Supercomputing Applications, the Pittsburgh Supercomputing Center, the University of Amsterdam, and the Masaryk University in the Czech Republic. Support was provided by the Albert Einstein Institute in Germany, the Poznan Supercomputing and Networking Center in Poland, DOE's Energy Sciences Network (ESnet), Sandia National Laboratories, SysKconnect, Hewlett-Packard and Force10 Networks Inc.

The numerical relativity group at the AEI, like many other groups working in computational physics, have a number of different projects involving different worldwide collaborations, and make use of a many different resources. Their [Cactus](#) black hole simulations are often very large, and expensive in terms of the resources they use, and it is important not to waste resources on simulations which are "going wrong". The results are very valuable, and are shared and studied by many people in the group.

Running their black hole simulations typically means submitting a job to a queue, where the simulation itself may only start after an hour, or a day, or a week. A persistent problem is how to keep track of, and manage, all the simulations queued by different people, and how to interact with their important results and visualize the data which they are waiting for.

Cactus simulations can be monitored, steered and visualized from different remote interfaces, such as from a normal [web browser](#). To be able to use these interfaces the researchers need to know when the simulations have started, and how to contact them.



The [Cactus portal](#), being developed by the [GridLab project](#), contains mechanisms for notifying users, and is starting to include interfaces for interacting with simulations.

Every Cactus simulation run by the AEI group automatically announces itself to the Portal (by including one thorn `Announce` in the simulation executable).

The simulation conveys information such as

- The machine where it is running.
- The local username under whose account it is running.
- The name of the parameter file.
- The name of the executable.
- The location of the data being created.

Periodically, the simulation updates the Portal with it's status, saying what iteration it is at, and providing an estimate of how much longer it will run for.

Users can log into the Portal and see a list of all the active (and old) simulations started by themselves, and by other members of their research group. They can see when the simulations started, which iteration they are currently at, and are provided with links to activate the simulations web browser steering interface, which provides also provides access to the data for visualization. Completed simulations are also listed, and provide a link which activates a remote file browser for viewing the simulation data.

The Portal not only provides a place for users to view the status of their simulations, it can also notify them of their status via Email or SMS messages. Users who provide their email and mobile phone numbers as part of their user profile can choose to receive status messages and updates about simulations. These interfaces are being extended so that users can choose the type of events they want to be notified about (e.g. only critical events) and also so that users can choose to be notified about simulations from other members of the group.

Portal uses underlying GRMS for submitting and controlling the user's jobs. Other services take part in the process. They include GridLab's Adaptive Components, monitoring services, information systems, data management services, security.

4. Summary

The GridLab meeting in Zakopane, the first Annual GridLab workshop was very successful. The GridLab project is very big in terms of number of institutions involved, number of workpackages and number of technical issues and problems on the way towards the integration of the overall software. Without having such a meeting we do not think the success of the project would be possible. That's why we have decided to have two such meetings a year. And always these meetings will be open to public in a small, constrained portion only, usually the first two days. The rest will be used to technical work, integration of the software and planning for the future developments. The meeting in Zakopane has been driving our activities for the next months and between this and the next meetings we will be working at our sites to improve the software and test it on the GridLab testbed. But, such a meeting not only allowed us to integrate the software but also people, which is very important in such multi-culture projects.