

CERN shares Grid expertise with EU projects

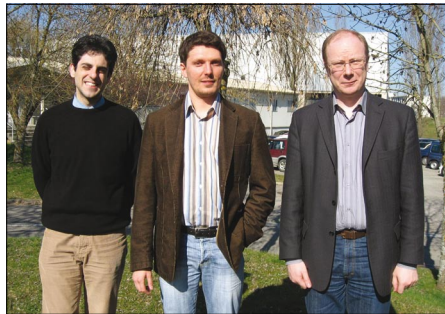
The European Union is funding a number of projects related to Enabling Grids for E-science (EGEE) to increase awareness of the Grid, develop applications and establish pilot infrastructures in new regions. CERN is a partner in several of these projects, all based on EGEE middleware (gLite). Associates from the different projects work within the IT department's Grid Deployment group (IT/GD), and at the end of their stay return to their home countries with an enhanced knowledge of the Grid.

Health-e-Child (www.health-e-child.org)

The Health-e-Child project is led by Siemens. It aims to build a scalable tool for clinicians who specialize in children's diseases, enabling them to share data among paediatric institutions in different countries. Sharing patients' data places security and confidentiality constraints on the software, and Grid authentication and authorization services are extended in a custom-built Grid gateway. The gateway also uses gLite data management services, including AMGA, to operate on metadata and access storage elements.

Clinicians can manipulate data, browse thumbnails of medical images and use a DICOM viewer to display full images. The software will be used to support decision-making in the fields of brain tumours, cardiology and rheumatology. The heterogeneity of imaging devices, technologies, data and protocols is a major challenge for the project.

A private Grid is being built to connect participating hospitals via the internet. Having set up the prototype infrastructure, CERN is now investigating the use of virtualization to minimize the number



The project associates in IT/GD, from left to right: Domenico Vicinanza (EUMEDGRID), Konstantin Skaburskas (Health-e-Child) and Rolandas Naujikas (BalticGrid).

of servers required at each institution. Working closely with the gateway development team at CERN has proved vital for efficient knowledge transfer and rapid prototyping. The project will run for four years and has just successfully completed its first-year review.

BalticGrid (www.balticgrid.eu)

The BalticGrid project aims to extend the European Grid by integrating new partners from the Baltic States (Lithuania, Latvia and Estonia). The Baltic countries are very advanced in terms of computing infrastructure – as was witnessed by Estonia being the first country to use e-voting for parliamentary elections. Although the project is coordinated by the Royal Institute of Technology (KTH) in Stockholm, having a competent project associate at CERN ensures that there is a close collaboration with EGEE in terms of operations and infrastructure.

CERN's main contribution to BalticGrid

is in SA1 (Grid operations), providing technical support for the installation, maintenance and operation of the BalticGrid production and certification infrastructures. This includes support and deployment of services such as Virtual Organization Membership Service, Service Availability Monitoring, Resource Broker, Workload Management System, LCG File Catalog, MyProxy and Information Indexes.

EUMEDGRID (www.eumedgrid.eu)

This project, which has received €1.6 m in EU funding, seeks to create national Grid initiatives in the Mediterranean area. EUMEDGRID has already established a human network and Grid presence in most countries in the Mediterranean basin, including Palestine and Israel. It hopes to add Syria to the list soon, showing that the Grid can extend scientific cooperation despite difficult political conditions.

One of the "gridified" applications demonstrated by CERN during the first EU review was a compute-intensive application developed by Italian and Tunisian scientists for modelling the intrusion of saltwater in coastal aquifers. This is useful for evaluating which pumping rates are sustainable without inducing contamination by saltwater. The project was a specific example of interdisciplinary collaboration between north and south Mediterranean countries.

CERN is not contributing hardware to EUMEDGRID but provides knowledge, expertise and support. At the end of the project it is hoped that sustainable Grid resource centres will operate in all of the participating countries.

John Shade, IT/GD

Editor's note

The article "CERN gLite community uses virtualization", published in the last issue of *CNL* (Vol. 42, issue 1, January–March 2007), describes some results based on several tools, including Xen and SmartDomains.

The following article expands on these two tools. Xen is a virtualization tool for supporting virtual servers, and SmartDomains is a virtual domain manager for Xen. SmartDomains has been developed by CERN openlab in collaboration with HP Labs to automate the deployment and management of virtual machines (VMs).

The vGrid tool described in the previous article is a web interface that makes SmartDomains easy to use in the context described in that article.

The CNL editors

SmartDomains deploys Xen VMs

What is Xen? Xen is a virtualization tool that enables a server to support multiple virtual servers. With Xen, users can share a server, isolate applications and boot different operating systems, saving time and money.

"This old box of mine is not really what I need, but for many reasons I cannot switch to a better OS, and by the way I do not need one but five machines, some with SLC to run gLite, some with Debian or Fedora; a small one that is always up for my personal web server, and two big ones that have a lot of memory and Geant4 libraries that would work overnights and weekends." Is this you? What would you say if booting or updating a pool of virtual machines (VMs) was as simple as pressing a button?

Xen is not pure magic – the software,

from XenSource and Cambridge University, has to be installed. This involves downloading a distribution that works, installing the files under /boot, updating /etc/grub.conf, and rebooting on Xen kernel. Xen is used more and more at CERN so help on this point is not hard to find.

Is it worth it? With Xen there is no need to buy more boxes, but still, if it takes until lunchtime to find the images, remember the configuration that worked, create logical volumes then launch several machines, there is little time left to actually use them before saving and compressing the images or cleaning everything. As soon as #!/bin/sh comes to mind it may be worth considering a switch to SmartDomains.

So what is SmartDomains? This is a virtual domain manager for Xen that is

powered by SmartFrog, an open source configuration management system developed at HP Labs. SmartDomains enables users to write configurations of VM pools and their workflows. It automatically prepares file systems and storage backends, boots VMs and checks their liveness, and finally cleans physical hosts back to their initial states, potentially saving and tarring images.

VMs can be launched in a whole computing centre without extra work. The description of the distributed virtual pool is kept in a local file that is used to launch the set-up with a simple start command on localhost. Finally everything can be cleaned in the same way. SmartDomains is based on a peer-to-peer layer that distributes the management work seamlessly.

With SmartDomains there is a language to learn, so those who are ruthless with computers may want to try a few manual VM boots before switching. For the others, starting with SmartDomains default configurations will simplify things.

SmartDomains was developed at CERN openlab by Xavier Gréhant (openlab fellow) and Olivier Pernet (openlab summer student 2006). Several people at CERN have experience in SmartDomains: we are

collaborating with the IT department's Grid Deployment group (IT/GD), which is using it in the context of gLite certification. Andreas Unterkircher is driving this use of the application. The description language used in SmartDomains enables template descriptions to be kept and extended. gLite testers edit their own templates via a web interface (vGrid), which was created by Omer Khalid to simplify the use of SmartDomains.

SmartDomains builds on SmartFrog's elaborate description language and its mechanism for deploying distributed Java components. As part of the openlab activities we are investigating applications for SmartFrog at CERN, especially in virtual resource management.

Since SmartFrog is an extensible component-based framework, SmartDomains is also extensible. With its complete base components for Xen it seems to have the potential to become a fully featured virtual resource management system; and because SmartDomains shifts the complexity of run-time resource management to the descriptions, it seems to be particularly adapted to launching batch execution environments. We are keen to find groups who are interested in developing use cases in these areas.

Support for virtualization

Within the IT department, Grid virtualization is organized as a stack of components that are developed or supported by the openlab team and the IT/GD group.

- Xen: support and expertise is provided by Havard Bjerke;
- SmartFrog: support is provided by Xavier Gréhant;
- SmartDomains: development and support is provided by Xavier Gréhant;
- vGrid: development and support is provided by Omer Khalid.
- Virtual machine converter: Dimitar Shiyachki contributed to its development.

Further information

For an online tutorial on SmartDomains, see <http://smartdomains.sourceforge.net>. To download SmartDomains, visit <http://sourceforge.net/projects/smartdomains>. For more information, contact xavier.grehant@cern.ch.

A presentation on the gLite test-bed can be found at twiki.cern.ch/twiki/bin/view/Virtualization/GDVirtualization.

See also <http://cern.ch/openlab> for news about openlab's activities.

Xavier Grehant, IT/DI, openlab

Parallel evolution: DILIGENT and ETICS

DILIGENT and ETICS, two European Commission-funded projects that are leaders in the European Grid scene, recently achieved milestone releases of their software, thanks largely to their collaboration with each other.

ETICS (E-infrastructure for Testing Integration and Configuration of Software) offers Grid-enabled automated building and testing of software. Just over a year after the project launched [in January 2006], its final release candidate was made public. The service, however, has been in use for some time by several projects, including DILIGENT (a Digital Library Infrastructure on Grid Enabled Technologies).

"We started to use ETICS when it was three months old for our build and deployment testing activities," explained DILIGENT's Pedro Andrade. "At the moment we are the only project also exploring some of the ETICS test functionality."

DILIGENT offers digital library management services that enable the dynamic creation and deployment of digital libraries.

"The way they are using the system gave us many requirements for the testing system," explained Alberto Di Meglio, the ETICS project manager. "For example, consider the case where you want to execute a test and have the system build or deploy components on a given node before the test starts. A number of specific



The ETICS interface and DILIGENT tools.

requirements for this case were written for us by DILIGENT."

Adopting software still in development for mission-critical tasks is a risky decision, but as Andrade explains, the collaboration has been successful.

"DILIGENT is a complex system, with software developed by seven partners and integrating 242 different software components, and ETICS also has a learning curve. It took some time but with the help of the ETICS team this was possible, and I think now, with our release alpha, we can say that the effort invested in the beginning has completely paid off."

The close collaboration between these projects is also helped by a common partner: Italy's Engineering Ingegneria Informatica SpA. Paolo Fabiani works on both projects for Engineering, writing the ETICS web application and managing the

building of the DILIGENT software.

"The collaboration has been very good," said Fabiani, "DILIGENT had some constraints and advanced requirements but ETICS was able to take these into account. Engineering is also interested in using ETICS for some pilot projects within the company, and we're also looking at ETICS for some EU projects starting later this year."

DILIGENT is planning the beta release of its software this summer, which will add a greater number of services to the basic package of the alpha release.

Meanwhile, ETICS has some 15 projects registered to use their service, including national efforts such as Grid-Ireland and the Italian EGRID. They are also involved with the US Grid community. The University of Wisconsin-Madison, a project partner, is proposing the ETICS service to the major US Grid projects like Open Science Grid and TeraGrid. ETICS is in the process of making plans for its second year, moving toward a 1.0 release of its service this spring.

Owen Appleton, EGEE contributing editor to International Science Grid This Week (iSGTW)

This article was published in full in *iSGTW* on 7 March (see www.isgtw.org/?pid=1000321). For further information see the DILIGENT website at www.diligentproject.org, and the ETICS website at <http://etics.web.cern.ch/etics>.