Enabling Grids for E-sciencE

Industry Forum Newsletter

Number 4 –October 2005

Table of contents

Editorial p.1

About Grid

- DEISA perspectives p.2

EGEE technical overview

- Security and biomedical applications: example of bioinformatic data encryption on the EGEE Grid Infrastructure p.3

Applications running on EGEE

- EGRID: a Finance Pilot Application for EGEE p.5
- Successful results for the grid Drug Discovery Application p.6
- gMOD: a Grid Application for the World of Entertainment p.7
- Prototyping Grid Analysis for Physics with ARDA p.8

News

- News from the Industry Forum p.9
- News from members p.9
- Upcoming Grid events p.10

Editorial

This editorial is the opportunity to provide an update of the status of application deployment on EGEE infrastructure, eighteen months after the project start-up.
In 2005 the scale of use of the EGEE production infrastructure by LHC experiments has increased significantly, reaching hundred thousands of jobs per experiment. As real production running by the LHC experiments is foreseen to generate raw data in the range 100 MB/s to 1 GB/s, service challenges were deployed to test the effective networking and mass storage capabilities. The first results for the generic capabilities have been very encouraging. In a 10 day period, an integrated rate of 600 MB/s was maintained from CERN to 7 national resource centres, namely FNAL, BNL, NIKHEF/SARA, FZK (Karlsruhe), INFN, IN2P3 and RAL. A total of 500 TB of data was transferred.

During July and August 2005, the biomedical community performed its first Data Challenge. Entitled WISDOM (Wide In Silico Docking On Malaria), the challenge saw over 46 million docked ligands during that period – the equivalent of 80 years on a single PC. Usually in silico docking is carried out on classical computer clusters resulting in around 100,000 docked ligands. This type of scientific challenge would not be possible without the grid infrastructure - 1000 computers were simultaneously used in 15 countries around the world. It also demonstrated the large scale usage of commercial docking software with floating licenses on the EGEE infrastructure. A total amount of 3000 floating licenses were given by BioSolveIT to SCAI, Fraunhofer, Germany for distribution on the grid. Maximum number of used licenses was 1008 on a particular day during the Data Challenge.

In 2005, EGEE Generic Applications have also significantly progressed and successfully moved from experimentation to real work. Working together as EGEE supported applications, different communities have come close to each other and started fruitful collaborative work. This has been true mainly within the Earth Sciences Research Virtual Organization and the MAGIC and PLANCK astrophysics communities. In fact, besides the provision of huge computational and storage

resources in Europe and world wide, one of the biggest merits of EGEE is to act as a catalyst of inter/multi-disciplinarity work which is at the basis of modern e-Science.

Since September 2005, in addition to the existing LCG-2 production infrastructure, a Pre-Production System has been set up by the SA1 team. The PPS complements the available gLite development prototype infrastructure maintained by JRA1 and available since the very beginning of the project which is mainly used to test new features, while the PPS is intended to allow stability and performance tests: resources and stability are the key ingredients for larger user involvement.

The feedback from the migration of several existing applications by the HEP, biomedical and generic teams to the new middleware has been an important input for the finalisation of the first version of the gLite middleware and its subsequent evolution. All this activity has also the essential role to prepare the scientific communities for acceptance of the new middleware, verifying improvements, and allowing time for adapting the applications and preparing the users to use the new infrastructure.

Finally, a very important indicator is that the overall number of users has approximately doubled in the last nine months. Most of the Virtual Organizations have demonstrated a very significant growth. There are some 60 Virtual Organizations which currently use the EGEE infrastructure. Many of these are small and of a regional nature, but some are well established international collaborations with more than 100 users.

Beside these success stories, there are important issues which applications are facing at the moment. Among them are the migration to gLite, the availability of gLite software on the pre-production and production services, the amount of effort which is still required by the Virtual organizations in order to obtain useful results from the use of the grid when compared to alternatives such as clusters and the availability and quality of user documentation. The good news is however that progress is being made in all of these areas as problems are being identified and being solved.

Vincent Breton, CNRS-IN2P3

About Grid

DEISA perspectives

After one and a half years of operation, the DEISA research infrastructure has reached production status in several of its planned services. The main objective of DEISA is to enable scientific discovery across a broad spectrum of science and technology, by the deployment and operation of a world class, persistent, production quality, distributed supercomputing environment. This becomes possible through a deep integration of existing national high-end platforms, tightly coupled by a dedicated network and supported by innovative system and grid software.

The DEISA supercomputing infrastructure includes a number of similar IBM AIX platforms glued together to create a “distributed virtual supercomputer”. The resulting platform is a super-cluster of computing nodes which are located in a few places in different countries, but which appears to end users as a single unified system. Several IBM supercomputers in Germany (FZJ and RZG), France (IDRIS), Italy (CINECA) and Finland (CSC) have been integrated in this way, and ECMWF is expected to join this distributed platform in 2006. The key software integration technology (in addition to the network itself) is the high performance network file system (in this case, IBM’s Global Parallel File System, GPFS) that enables transparent and efficient access to remote data in the distributed platform. This platform has reached operational status, and GPFS has confirmed its remarkable performance: remote file accesses employ all the available network bandwidth. The high added value of this integrated platform arises from the possibility of redistributing the computational workload by migrating jobs across national borders, in order to free huge resources for one specific application in one site.

Three leading supercomputing centres (BSC in Barcelona, HLRS in Stuttgart and LRZ in Munich) have been integrated as new partners

from May 1st, 2005. They are operating other different supercomputing platforms, and this has substantially increased the relevance of the design and deployment of a heterogeneous supercomputing Grid. Three basic services are planned in the heterogeneous environment (which includes, of course, the IBM systems of the AIX super-cluster), namely, support for workflow applications, deployment of a global data management environment that would enable access to distribute data sets, and co-allocation on the extended heterogeneous Grid. The first of these services is already implemented using UNICORE, the others will be deployed in 2006.

Last but not least, DEISA has moved aggressively into the applications domain. The DEISA Extreme Computing Initiative was launched in April 2005. A European call for proposals for demanding, grand challenge simulations using the resources of the infrastructure was launched, and 53 proposals were received. After evaluation by national committees, 27 of these proposals were retained for operation in 2005-2006. The typical profile of these applications is either multi-site workflow applications or single site applications requiring exceptional computing resources (up to 1,4 million hours have been allocated to a single application). In order to facilitate user adoption of the infrastructure, an Applications Task Force of applications architects has been created to help users design, adapt and optimize applications to the DEISA infrastructure. Further details on scientific projects are available at the DEISA Web (www.deisa.org). The list and abstract of all the projects in operation in 2005-2006 will be available on our Web site after November 8, 2005.

Victor Alessandrini, IDRIS – CNRS
(va@idris.fr)

EGEE technical overview

Security and biomedical applications: example of bioinformatic data encryption on the EGEE Grid Infrastructure

Biomedical Security Requirements

Biomedical applications work on different data sets such as proteomes, genomes and medical images. This data in the bioinformatic or medical areas can concern patients, and can be used in scientific or industrial process such as drug design or gene function identification. Then this data need to have a certain level of confidentiality and integrity to preserve the patient privacy or the patent secret. And these requirements are important in the context of a Grid such EGEE, where the computing and storage resources are distributed across the world. Biomedical applications are a pilot application area in the EGEE project and manage a devoted virtual organisation: the biomed VO (see http://egee-na4.ct.infn.it/biomed).

Biomedical science has specific security requirements such as electronic certificate systems, fine grain access to data, encrypted storage of data and anonymisation. Certificate system provide biomedical entities (like users, services or Web portals) with a secure and individual electronic certificate for authentication and authorisation management. One key quality of such a system is the capacity to renew and revoke these certificates across the whole grid. Biomedical applications also need fine grain access (with Access Control Lists, ACL) to the data stored on the grid: biologists and biochemists can then, for example, share data with colleagues working on the same project in other places. This data also need to be gridified with a high level of confidentiality because it can concern patients or sensitive scientific/industrial experiment. The solution is then to encrypt the data on the Grid storage resources, but to provide authorised users with transparent and unencrypted access.

EGEE Security Overview

The EGEE grid is composed of three different platforms:

- the production platform, in operation since the beginning of the project and originally using the LCG-2 middleware inherited from the previous European DataGRID project, but now also deploying a number of key gLite components
- the prototype platform used by the gLite middleware developers

- the pre-production service used to validated the gLite components by the application testing team.

The security components available today on the LCG-2 platform are, for example:

- authentication, based on X.509 certificates and Public Key Infrastructure (PKI), which permits users to authenticate themselves to access the EGEE Grid infrastucture,
- the ability to authenticate only once per session through a Single Sign On (SSO) mechanism,
- monitoring and logging services which permit the identification and the analysis of all actions on the platform and also to prevent attack or intrusion from crackers.

The gLite developers (JRA1 activity) and the security team (JRA3 activity) are currently working to extend the security services on the future EGEE grid platform with gLite middleware:

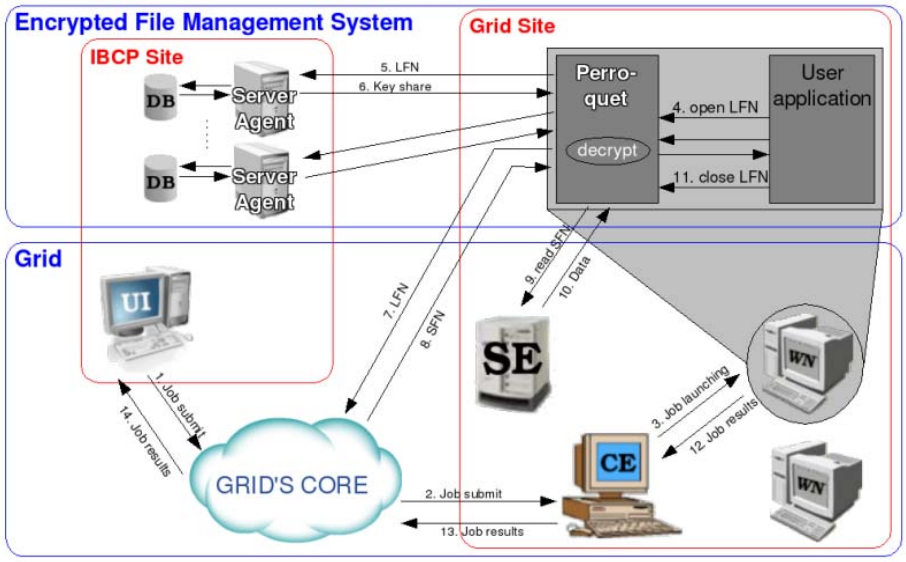
- data encryption for long term storage, confidentiality and integrity,
- fine grain access to data and resources access through ACL mechanisms and VO Membership Service (VOMS),
- isolation of user tasks when computing on the nodes of the grid in a local context with the Local Credential Mapping Service (LCMAPS).

Other components are planned for the future, such as:

- dynamic connectivity service to manage and authorize the connection between a user task on a grid node and the rest of platform,
- pseudonymity, to allow an user to access some grid components without giving its real identity

Application to GPS@ bioinformatic portal: Encrypted protein data on EGEE grid.

Bioinformatic analysis of data produced by complete genome sequencing projects is one of the major challenges of the coming years. Integrating up-to-date databanks and relevant algorithms is a clear requirement of such an analysis. Grid computing, such as the infrastructure provided by the EGEE project, would be a viable solution to distribute data, algorithms, computing and storage resources for Genomics. Providing biologists with a good interface to Grid infrastructure will also be a challenge that must be adressed. The GPS@ web portal, Grid Protein Sequence Analysis (http://gpsa.ibcp.fr), aims to be such a user-friendly interface for these genomic Grid resources on the EGEE infrastructure: an integrated Grid portal devoted to molecular bioinformatics. The current version is under development and deployed on the production LCG-2 based production service of EGEE. GPSA is a porting experiment of the NPS@ (Network Protein Sequence Analysis, http://npsa-pbil.ibcp.fr) services to the EGEE grid.



Architecture of the Perroquet encrypted data management system, and its integration into the EGEE grid platform and middleware (*UI: user interface, CE: computing element, SE: storage element, WN: worker node, DB: database, LFN: logical file name*)

NPS@ is a production Web portal hosting protein databases and algorithms for sequence analysis. Our server is computing more than 5.000 bioinformatic analyses per day, coming from numerous users and countries. Currently the submitted data is processed locally on the IBCP resources and not exposed to the rest of the world. Deploying these jobs and data on a distributed environment, as we are doing on the EGEE grid through the GPS@ porting experiment, will bring this data on remote

computers for storage and job computing. This provides a security risk as these remote computers can be compromised, especially as there are involved in a distributed environment.
In this context, our bioinformatic algorithm should be processed on an encrypted version of the user data, but this bioinformatics software should be considered as a "black box" (source code is not always available, and there are often too many applications to think of modifying them) and they need standard input/output access to data, *i.e.* no remote access to the storage element (SE) of the EGEE grid. Consequently, accesses to encrypted gridified data must be transparent, as if it were local files. As we said, JRA1 and JRA3 activities of EGEE project are developing encrypted file management, but for gLite middleware, so on the pre-production platform only. Our bioinformatic portal GPS@ is working on the production platform, mainly using the LCG-2 middleware.
We are developing an encrypted file management system, "Perroquet" (see Figure) based on the Parrot software (http://www.cse.nd.edu/~ccl/software/parrot) that we have modified to made compliant with our bioinformatic applications (mainly standard input/output and unencryption on the fly) and with the EGEE production platform. The current prototype of Perroquet is used in our GPS@ Web portal to secure sensitive user data on the EGEE resources. Applications can encrypt data and transparently access the encrypted files with standard I/O. Files are decrypted on the fly: there is no decrypted data on the remote worker nodes. Perroquet also ensures the remote access to the data from the worker nodes (WN) to the storage element, through transfer protocols compliant with the LCG-2 middleware such as gsiftp, ftp or http. The cryptographic keys used to encrypt/decrypt biological data follow the AES standard (with key length of 256 bits) and are stored on several servers with a M-of-N technique. Keys are split in N shares, each one is stored in a different server, and the reconstruction of the key needs exactly M of these shares. Such a technique has good security properties because, with less than M shares, we cannot deduce more information about keys than if we have random data. So, an attacker, who would want to retrieve a key, will have to compromise at least M servers. The performance of the system is still reasonable despite the cryptographic overhead: the cost for manipulating encrypted data (unencryption overhead) was estimated over large files (200MB). It proves to be negligible for the target applications.

For further information on the Perroquet system or the GPSA bioinformatic portal, please contact the authors or see on http://pbil.ibcp.fr.

R. Mollon and C. Blanchet
(r.mollon & c.blanchet@ibcp.fr)
Institute of Biology and Chemistry of Proteins
CNRS, Lyon, France

EGEE Applications

EGRID: a Finance Pilot Application for EGEE

EGRID is a MIUR (Italian Ministry of University and Research) funded research project, originally setup to investigate the role of grid technologies in the field of Complex Systems applied to Economics and Finance. With the aim of better co-ordinating the efforts of three independent Italian research projects in these fields and giving raise to synergies among them, the EGRID project was also assigned a further task. EGRID was to implement a Grid system allowing geographically distributed scientific communities involved in Economics and Finance to share data as well as applications.

A first version of this system was released in October 2004, and is presently in operation. Technically, the infrastructure is based on LCG/EGEE middleware and EGRID acts as a virtual organization within the GRID.IT computational grid. Currently, 500GB of financial data are stored and made available to the user community, consisting of two partner projects: one headed by Prof. Mantegna from the University of Palermo, and involving seven other research institutions in Italy; and another one headed by Prof. Fanni from the University of Trieste, hosted at the Area Science Park of Trieste.

To ease data processing, specific Grid tools were developed, comprising both filtering procedures to extract subsets of data, and

simulation and modelling packages to run on the data, allowing the submission and management of hundreds of jobs. Privacy and security issues related to High Frequency Data coming from various stock exchanges proved to be challenging: in order to cope with it, EGRID is now actively collaborating with INFN-CNAF (the Italian National Institute for Nuclear Physics Center for Telematics and Informatics) and INFN-CERN in developing StoRM, an implementation of the Storage Resource Manager standard with enhanced features for data privacy. This tool will be incorporated in the second version of the system now under way.

The design of the second version of the system is based on the experience gained so far, and on the feed-back provided by our users. EGRID is now also extending this experience to other European partners, and for this reason it was recommended for acceptance by the EGEE Generic Application Advisory Panel during the third EGEE conference in Athens, and is pending approval for transfer to the EGEE production service.

The EGRID project is organizing the international workshop "Grid in Finance 2006" that will act as an international forum for researchers in academia and in industry, from the field of grid-based applications in Finance. The call-for-papers was released at the beginning of October and can be downloaded from the website of the conference: www.gridinfinance.org.

For further info about the project, please see the official website www.egrid.it or send an e-mail to info@egrid.it

"Grid in Finance 2006" Worshop
1st International Worshop on Grid Technology for Financial Modeling and Simulation.
Palazzo Steri Palermo, Italy, 3-4 February 2006.
See www.gridinfinance.org

Stefano Cozzini (cozzini@democritos.it)

Successful results for the grid Drug Discovery Application

The Drug Discovery Application has been running on the EGEE production service since December 2004, as a joint collaboration between the Laser Processing Consortium (LPC) and the Fraunhofer Institute for Algorithms and Scientific Computing (SCAI) in the WISDOM project (Wide *In Silico* Docking on On Malaria). The aim of this application, where scientists carry out *in silico* binding (or docking) experiments, is to find potential new drugs to combat Malaria, a dreadful disease which kills a million people per year and affects 300 million more, mostly in the third world. *In silico* docking enables researchers to compute the probability that potential drugs will dock on the active site of one of the malaria parasite proteins.

A first stress test was successfully achieved in December 2004 when – thanks to the Grid – only 2 days where required by the application to analyze 100 000 potential drugs, while 6 months would normally have been required on a single computer.

A couple of months later, a new challenge was successfully carried out. From 11 July to 19 August 2005, over 46 million molecules were analysed – the equivalent of 80 years on a single computer – thanks to a deployed on 1000 computers within the EGEE infrastructure, spread across 15 countries
This was allowed by the use of two different computer programs dedicated to prediction of protein binding. The first, FlexX, was provided free of charge by BioSolveIT (a spin-off of SCAI) in Germany, and used in July. The second program, AutoDock, a piece of free access software, was used in August and the results compared to those from FlexX.
1000 licenses were given by BiosolveIT for FlexX, an extremely fast, robust and highly configurable (FlexX-able) computer program, first commercialised in 1997. A central license server was set up at SCAI in order to manage all the license requests of the distributed computers in the BIOMED virtual organization. It is probably the first time that a software company has offered so many licenses for a public interest research program.

The researchers involved in the Drug Discovery Application are now assembling, sorting and filtering the huge quantity of results provided by the summer experiments. The first results are expected to be presented on the web at http://wisdom-eu-egee.fr at the end of October. The full information will be made available in a shared database.

The next step will consist of the experimental validation of the results by *in vitro* testing of the best identified molecules. This will be done, probably at the beginning of 2006, by the Center of Bio-Active Molecules Screening (CMBA), which constitutes the first French academic high-throughput screening platform dedicated to cellular systems.

Nicolas Jacq (jacq@clermont.in2p3.fr) and Marc Zimmermann (Marc.Zimmermann@scai.fraunhofer.de)

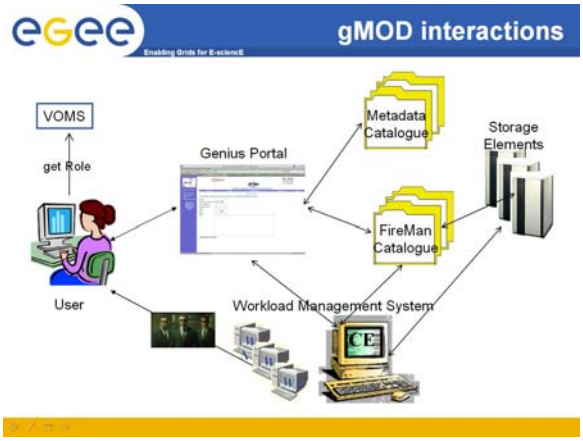
gMOD: a Grid Application for the World of Entertainment

The use case

The gMOD application (grid Movie on Demand) is a new and exciting system developed by the GILDA Team, to show up how the Grid can be used in the world of Entertainment.
gMOD is one of the services plugged into GENIUS (https://genius.ct.infn.it), the EGEE Grid portal jointly developed by INFN and NICE, which allows scientists to exploit the Grid using only a web browser.
The goal of gMOD is to offer users a Video-On-Demand service. They are presented with a list of movies (movie trailers in our case due to license issues) to choose from and once they have made a choice, the video file is streamed in real time to the video client in the user's workstation.
Users can also inspect the details of every movie listed such as Title, Runtime, Country, Release Date, Genre, Director, Cast and Plot Outline and moreover they can issue queries on one or more attributes to find specific movies fulfilling their criteria.

Under the hood

gMOD is built on top of the new EGEE gLite middleware and makes use of many gLite services: The FiReMan File Catalogue and Data Management System, the AMGA Metadata Catalogue, the Virtual Organization Membership Services (VOMS), Storage Elements, and the Workload Management System (WMS). Movie files are physically located at different Storage Elements sited in different places, but the File Catalogue keeps track of where they can be actually found. The Metadata Catalogue is the repository of the detailed information for each movie present in the File Catalogue. Queries are possible thanks to its presence. VOMS is used to establish and enable users' rights depending on their roles (normal users and managers). The Workload Management System does the hard work: it is responsible for retrieving the chosen movie from the right Storage Element and streaming it over the network down to the user's desktop or laptop.



Below it is a description of how the different services interact during a search for a movie:

- the user initialises his/her proxy certificate contacting a VOMS server that gives him/her the right access rights;
- a query on some attributes specified by the user is executed on the Metadata Catalogue;
- one of the movies satisfying the user's query's criteria is chosen by the user;
- the movie is searched against the File Catalogue, which returns the proper location (Storage Element) of the relative file;
- the WMS finally looks for an available Computing Element and assigns to it the

task of retrieving and streaming the movie directly to the user's workstation.

It is worth noting that gMOD has been designed with the commercial issues and technical problems of a Video On Demand service in mind, but can also be used to retrieve any kind of digital multimedia contents from the network, with many possible interesting applications such as e-Learning Systems and Digital Libraries.
If you want to experience gMOD yourself or to get more information about it, please contact the GILDA team (grid-prod@ct.infn.it) or the authors.

Tony Calanducci (tony.calanducci@ct.infn.it) and Giuseppe La Rocca (Giuseppe.larocca@ct.infn.it)

Prototyping Grid Analysis for Physics with ARDA

ARDA, a LCG project co-funded by EGEE

Within the Large Hadron Collider (LHC) Computing Grid (LCG) project -which aims is to build and maintain a data storage and analysis infrastructure for the entire community that will use the LHC- was created in 2004 the ARDA project (A Realisation of Distributed Analysis for LHC). The mission of ARDA is to develop a prototype Grid analysis system for the experiments at the LHC experiments.
ARDA actually started in parallel with the EGEE project -which aims to provide a dependable Grid infrastructure for a variety of users from different scientific domains- and was jointly funded by LCG and EGEE. The guiding principle behind ARDA is that exploring the opportunities and the problems encountered in using the Grid for LHC analysis would provide key inputs for the evolution of the EGEE gLite middleware.

In EGEE, the LHC experiments as well as a wide community of biomedicine applications have key roles in driving the evolution of the infrastructure. At the same time, they are major users of the computing infrastructure. The LHC experiments contribute to the Grid evolution through initiatives like ARDA, which provides the opportunity of exploring new ideas and prototyping advanced services not yet provided by the infrastructure, which might lead to the development of innovative applications.

A prototype activity inside ARDA

The patterns observed in the past suggest that the new Grid infrastructure will enable and stimulate new approaches to data handling and analysis for physics, with the ultimate goal of enabling a large scientific community to maximize the scientific output of the LHC programme. A sound approach to such an evolving infrastructure is to prototype the future systems together with the users, exposing them early to the Grid environment early on, and discussing the evolution on the basis of the experience gained.

Testing the Grid under real conditions also gives effective feedback to the developers of Grid middleware. During the first phase of EGEE, ARDA played a key role in the testing of the middleware, with its access to "previews" of gLite components. Progressively this activity moved towards detailed studies of performance issues, but always using the analysis scenario as a guideline. As an example, the experience and requirements of the LHC experiments led ARDA to propose a general interface for metadata access services. Eventually an ARDA prototype called AMGA (ARDA Metadata Grid Application) made its way into the gLite middleware and is now also used by non-HEP applications.

The LHC experiments welcomed the idea of a prototype activity agreed with ARDA. It was decided to propose to each LHC experiment an independent prototype activity, since it was considered unrealistic to force at an early stage commonality in the use of tools at an early stage, since each experiment has different physics goals and data organisation models. On the other hand, all different activities hosted in the ARDA team benefit from common experience and cross fertilisation.

We can give here two examples of prototypes activities inside ARDA.
The ARDA-LHCb prototype activity is focusing on the GANGA system (a joint ATLAS-LHCb project). The main idea behind GANGA is that the physicists should have a simple interface to their analysis programs. GANGA allows the preparation of the application, organisation of the submission and gathering of results via a clean Python API. The details needed to submit a job to the Grid (like special configuration files) are factorised out and applied transparently by the system. In other words, it is possible to set

up an application on a portable PC, then run some higher-statistics tests on a local facility (like LSF at CERN) and finally analyse all the available statistics on the Grid simply by changing the parameter that identifies the execution back-end.
The ARDA-CMS activity started with a comprehensive evaluation of gLite and existing CMS software components. Eventually ARDA focused on providing a full end-to-end prototype called ASAP, prototyping the Task Monitor and the Task Control services. The Task Monitor gathers information from different sources: MonaLisa (providing mainly run time information via the CMS C++ framework); the CMS production system; Grid-specific information (via the gLite R-GMA system). The Task Control implements CMS-specific strategies, making essential use of the Task Monitor information. Task Control understands the user tasks (normally a set of jobs) and organizes them in a way the user can delegate many tasks like job control and error recovery. Some key components of this very successful prototype, which incorporated a lot of feedback from real users, are now being migrated within the official CMS system CRAB in the framework of the CMS-LCG taskforce.

ARDA, a very active contributor to EGEE

ARDA is a direct and very active contributor to the EGEE project. EGEE benefits from the exchange of ideas and from the experience of groups that actively support scientific communities, notably the LHC experiments and biomedicine. Being able to map concepts and strategies developed in one particular application domain to other sciences is a powerful indicator of an improved theoretical understanding of Grid techniques and fully in line with EGEE's goals of supporting a truly multi-science user community. Each application domain can contribute with its specific requirements and knowledge and thus improve the system as a whole.
The different prototype activities in ARDA, together with other activities within the experiments, are converging on a first version of the distributed-analysis systems that will be used in the first phase of LHC operation. In the second phase of EGEE we hope to streamline this activity to further support the experiments' systems for both production and analysis. This means also continuing to influence the evolution of the middleware and the Grid infrastructure, using larger-scale experience and fostering the contacts with non-HEP scientists established during the first phase.

Massimo Lamanna, CERN/IT/GD
(Massimo.Lamanna@cern.ch)

News

News from the Industry Forum

Industry Forum meeting in Pisa

The next EGEE Industry Forum meeting will take place on Tuesday 25rst October 2005 3:00 PM in Pisa, Italy, during the fourth EGEE conference.
The general topic will be Grid and ISV (Independent Software Vendors).

Information about gLite and Gilda

The Industry Forum has worked on the realization of two flyers concerning the Gilda infrastructure (http://public.eu-egee.org/industry/ifdocuments/gilda-flyer.pdf) and gLite, the EGEE's next generation middleware (http://public.eu-egee.org/industry/ifdocuments/glite-flyer.pdf)

Industry Forum Website

The Industry Forum Website has been updated during the last two months. See http://public.eu-egee.org/industry/

Grid Web-based tutorials

The Industry Forum would like to organise quarterly Grid web-based training for its members.
Each training session would consist of web-based material (i.e. on-line tutorials) in parallel with two conference calls of approximately 2 hours each, separated by one week.
In order to help us better target these training sessions, could you please take a few minutes to fill in this quick survey.

- Would you be interested in such trainings? Why?
- What topics would you be interested in?
- Does the duration and periodicity seem appropriate? Why?

Any other comment is welcome.

Please e-mail your answers to Myriam Brun, brun@mas.ecp.fr indicating name, position, email, phone, company name, company area, and web site.

News from members

ATOS ORIGIN's vision on GRID technology and the NESSI initiative

Over the last few years, Grid technology has evolved from a technology mainly designed for the needs of the High Performance Computing (HPC) community towards an open framework, often referred as Next Generation Grid (NGG), supporting the business domain in general. This development from Science Grids to Business Grids has been complemented by efforts to create the Knowledge Grid by integrating the achievements from the Semantic Web community. These efforts aim to reduce the complexity of Grid-enabled systems and a complete virtualization of Grid resources and ATOS ORIGIN is participating in this research as an active actor according to our vision and the high expectations we have for this technology.

There have been a significant number of middleware solutions and interesting systems resulting from EU Grid research programmes, National initiatives and other on-going EU activities. However, the dream of global grid computing has not caught on with most business, and the promise of delivering boundless supercomputer power to any user and anywhere still seems far away. There is a clear threat that industry and companies do not accept these middleware solutions at large simply due to lack of knowledge or lack of confidence in their maturity and reliability. For example, we can look at the well known Globus Toolkit, in which many sponsors companies like IBM, HP, Fujitsu, Intel and others put high expectations. Despite what many people would like to read, it is still not seen as a solid product for business and we regret to say that this is the current perception about Globus Toolkit from many of our customers. Globus, in its 4th version (GT4), is incorporating many improvements for developers, and although vendors are saying that they will embed some parts of it in its products (SAP solutions, IBM WebSphere, etc) this IT uptake may be slow.

ATOS ORIGIN believes that today, the Grid is at a critical phase in its transition from research and academic use to broad adoption. We have defined a clear strategy round the adoption of this promising technology for our customers whose first priority is still to reduce the TCO (Total Cost of Ownership) of computing resources and optimise their internal processes.

A lack of awareness of the benefits brought by the use of Grid technologies and the lack of reference business cases (to persuade potential users) is leading to a weak commercial exploitation of results and so of the general deployment of this technology into the market, missing the EU´s competitiveness and leadership in this technological area.

ATOS ORIGIN considers that it is time to establish effective routes to push this technology adoption and to stimulate the research into innovative business models and early Grid pilots, since this momentum fits with the consolidation of some implementation of WSRF (Web Service resource Framework) specifications.

ATOS ORIGIN is founder member of industrial NESSI initiative (Networked European Software & Services, http://www.nessi-europe.com) whose major objective is to develop a visionary strategy for software and services driven by a common European Research Agenda and hence transforming the European economy into a knowledge based economy and enabling the European software and IT services industry to attain a stronger global position.

Josep Martrat and Santi Ristol, ©ATOS ORIGIN, 2005

Upcoming Grid events

SC 2005 (featuring The 6th IEEE/ACM International Workshop on Grid Computing)

Location: Seattle, United States
Date: 12-18 Nov 2005

SuperComputing 2005 (SC05), the international conference on high performance computing, networking and storage, will convene in November 2005 in Seattle. Under the theme, "Gateway to Discovery," SC05 will showcase how high performance computing, networking, storage and analysis lead to advances in research, education and commerce.

For more information please visit:
http://sc05.supercomputing.org/

Grid 2005 - 6th IEEE/ACM International Workshop on Grid Computing

Location: Seattle, United States
Date:14 Nov 2005
Grid 2005 is an international meeting that brings together a community of researchers, developers, practitioners, and users involved with the Grid. The objective of Grid 2005 is to serve as a forum to present current and emerging work as well as to exchange research ideas.
For more information please visit:
http://pat.jpl.nasa.gov/public/grid2005/

Internet Users Conference - CFP

Location: Dubrovnik, Croatia
Date: 21-23 Nov 2005
The Croatian Academic and Research Network is inviting you to actively participate in the Internet Users Conference CUC 2005, contributing with your paper or presentation.
For more information please visit:
http://www.carnet.hr/cuc

IEEE Conference and WS on e-Science and Grid Technologies

Location: Melbourne, Australia
Date: 5-8 Dec 2005
Grid research has focused on the marshalling of computation and data, and their interactions at hubs of analysis and synthesis. This focus has spawned the notion of computational and data Grids, respectively. As the technology for these has matured, however, increasing attention is being directed towards the actual sources of data: the instruments and sensors.
For more information please visit:
http://www.gridbus.org/escience

Grid in Finance 2006 Worshop

Location: Palazzo Steri Palermo, Italy
Date: 3-4 Feb 2006
1st International Worshop on Grid Technology for Financial Modeling and Simulation.
For more information please visit www.gridinfinance.org

The First International Conference on Grid and Pervasive Computing

Location: Tunghai University, Taichung, Taiwan
Date: 3-5 May 2006
The International Conference on Grid and Pervasive Computing (GPC) is an annual international conference on the emerging areas of merging grid and pervasive computing, aimed at providing an exciting platform and paradigm for all the time, everywhere services. Grid and Pervasive Computing (GPC) is a forum for scientists, engineers and practitioners throughout the world to exchange ideas and research results related to the design, use, analysis and application in the field of grid computing and pervasive computing.
For more information please visit:
http://hpc.csie.thu.edu.tw/gpc2006

Anyone interested in joining the EGEE Industry Forum should contact either Christian Saguez (christian.saguez@ecp.fr), Guy Wormser (wormser@lal.in2p3.fr) or Myriam Brun (brun@mas.ecp.fr)

Enabling Grids for E-sciencE