

## APPLICATIONS RUNNING ON EGEE

The Enabling Grids for E-scienceE (EGEE) project initially focused on two well-defined application areas, Particle Physics and Life Sciences, due to the fact that these communities were already Grid aware and ready to deploy challenging real applications at the beginning of the project.

EGEE has since established a broader portfolio of applications across a wide range of sectors, including Earth Sciences, Astroparticle physics and Computational Chemistry. At the last project conference in November 2004, the EGEE Generic Applications Advisory Panel (EGAAP) recommended four new applications to join the EGEE Grid infrastructure: Cosmology (Planck project), Drug discovery (the search for new drugs against malaria), the extension of Earth Sciences (Hydrology) and the GRACE project (GRid enabled seArch and Categorization Engine).

### High Energy Physics (HEP) Applications

Since the start of the project, the four particle physics Large Hadron Collider (LHC) experiments **ALICE** (A Large Ion Collider Experiment), **ATLAS**, **CMS** (the Compact Muon Solenoid experiment) and **LHC** (The Large Hadron Collider Project), have made large scale use of the EGEE infrastructure. These projects will continue to do so in preparation for the switch-on of the Large Hadron Collider at CERN in 2007. Each experiment has different physics goals but all need to perform massive simulation studies of the 'events' that will be produced when high energy beams of protons or heavy ions collide. An example of this type of usage has been that of the Large Hadron Collider beauty experiment (LHCb) which generated some 300 million events using resources distributed world-wide (63 sites). At one point there were up to 4,500 PCs running LHCb jobs. They have generated more than 100 Terabytes of data distributed in centres throughout Europe.

Each LHC experiment has operated in a multi-grid environment resulting from their individual history and constraints. The two largest experiments, ATLAS and CMS, have a significant membership in the United States included in their total of ~200 institutes. Each has used the GRID3 project in addition to EGEE, and provided operational interfaces to these grids. ATLAS also makes use of NorduGrid. Both ALICE and LHCb developed their own distributed grid-like technology before EGEE in the ALIEN and DIRAC systems, and indeed some of the technology used in the systems for data and job management is being used to drive development of the new gLite software.

Other international HEP experiments have made use of the EGEE infrastructure, including **Babar** (the B and B-bar experiment), **CDF** (Collider Detector at Fermilab) and **DØ** experiments based on United States colliders, and the **ZEUS** experiment based on the collider in Germany.

The **ARDA** project (A Realisation of Distributed Analysis for LHC) has developed prototype analysis systems interfacing with the gLite middleware which will enable the HEP experiments to move to individual user analyses. This type of usage is de-centralised and comes from hundreds of individuals, rather than tens of production managers.

## Biomedical Applications

Applications in the Biomedical arena have been included in the EGEE project from the outset. Nine regularly used applications are now available and several new initiatives are currently being considered.

The applications available include: **GATE** (a radiotherapy planning and medical tomography application); **GPS@** (Grid genomic web portal) a portal to various gridified protein analysis tools; **CDSS** (Clinical Decision Support System) which aims at extracting medically relevant knowledge from a large set of information with the objective of guiding the practitioner in their clinical practice; **gPTM3D** (interactive radiological image visualization and processing tool) which addresses the problem of running interactive tasks on a grid for supervised medical data exploration and analysis); **Mammogrid** which aims to develop a European-wide database of mammograms that can be used to investigate a set of important healthcare applications as well as potentially support effective co-working between healthcare professionals throughout the EU; **SIMRI 3D** is an image simulator service for the research community to better understand the complex Magnetic Resonance (MR) technology; **xmipp\_MLrefine** (Macromolecular 3D structure analysis) focuses on 3D structure analysis of macromolecules from (very noisy) electron microscopy images; **GridGRAMM** (Molecular Docking web): is a simple interface to do molecular docking on the web. It can currently generate matches between molecules at low or high resolution, both for protein-protein and ligand-receptor pairs; **GROCK** (Mass screenings of molecular interactions web) aims at providing an easy way to conduct mass screenings of molecular interactions using the web.

Three new biomedical applications are currently being ported and tested on the LCG2 infrastructures in the fields of genomics and drug discovery:

**SPLATCHE** (genome evolution modelling) is a modelling tool which simulates the spread of humans into geographically realistic landscapes, taking into account the spatial and temporal heterogeneity of the environment and generating the molecular diversity of one or several samples of genes drawn; Docking platform for tropical diseases: this **Drug Discovery** application aims to find potential new drugs against malaria - the world's second biggest killer, after tuberculosis. In silico docking enables researchers to compute the probability for a given chemical compound to dock on a target active protein; **Xmipp\_assign\_multiple\_CTFs** (Micrographia CTF calculation) is a tool for determining the Contrast Transfer Function impairing images acquired by electron microscopes.

## Generic Applications

The **GRACE** (GRid enabled seArch and Categorization Engine) project ended in February 2005. The aim of the project was to implement a distributed search and categorization engine based on grid technology. GRACE provided concrete and early feedback to EGEE on the usage of the grid and contributed to the dissemination of grid awareness.

The **MAGIC** application simulates the behaviour of air showers in the atmosphere, originated by high energetic primary cosmic rays. These simulations are needed to analyse the data of the MAGIC telescope, located at the Canary Islands, to study the origin and the properties of high energy gamma rays. The first data challenge started on EGEE infrastructure began at the beginning of March 2005.

The ESA **Planck** mission aims to map the microwave sky performing at least two complete sky surveys with an unprecedented combination of sky and frequency coverage, accuracy, stability and sensitivity. The satellite will be launched in 2007 carrying a payload composed of a number of microwave and sub-millimetre detectors which are grouped into a high frequency instrument (HFI) and a low frequency instrument (LFI) covering frequency channels ranging from 30 up to 900 GHz.

There are five different domains within Earth Science Research: Geoscience, Hydrology, Earth Observation, Climate and Solid Earth Physics. In each domain at least one application has been deployed on EGEE and GILDA.

Geocluster, the leading industry Seismic Processing Solution, is the first industrial application successfully running on the EGEE Grid Production Service, within the Expanding GEOsciences on DEMand (EGEODE) Virtual Organisation (VO). EGEODE is dedicated to research in **Geosciences** for both public and private Industrial Research and Development and Academic Laboratories. It enables researchers to process seismic data and to explore the composition of the Earth's layers.

**Earth Observation** has deployed the GOME satellite ozone profiles as well as data from the ERS/SAR satellite experiment used for oil spill detection. The grid will assist scientists to analyse ozone retrieval by enabling a single computing environment for the different steps, allowing the data to be easily shared between "producer" and "consumer".

**Solid Earth Physics** applications deal with earthquake mechanisms and numerical simulation of earthquakes in complex 3D geological models.

The first **Hydrology** application investigates the impact of exploitation uncertainty on seawater intrusion in a coastal aquifer of the Mediterranean basin, using Monte-Carlo simulations based on a 3D density-dependent groundwater flow and salt transport model.

In relation to **Climate**, a flood forecasting application consisting of a cascade of simulations initialised with experimental data and using meteorology, hydrology and hydraulic models, is being migrated from the CrossGrid test bed to EGEE.

For more information about the Applications running on EGEE, visit the NA4 webpage's at: <http://egee-na4.ct.infn.it/index.php> or contact Florence Jacq, email: [florence.jacq@clermont.in2p3.fr](mailto:florence.jacq@clermont.in2p3.fr) .