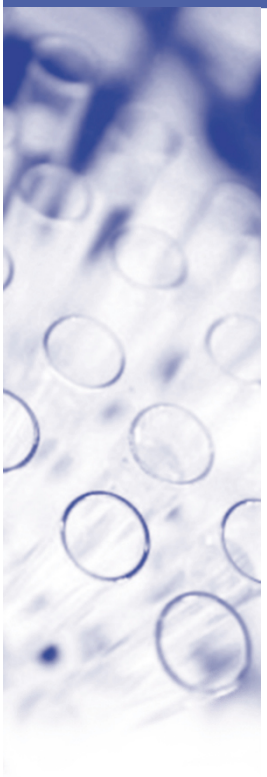

BIOLOGICAL APPLICATIONS

Introduction to Enabling Grids for E-science in Europe (EGEE)

EGEE is a European initiative (with links worldwide) to establish a stable, reliable, production quality Grid infrastructure for e-Science within the next two years. It provides support for porting applications to the Grid and dedicated training at all levels. EGEE expects to have more than 3000 users in more than 15 countries in the first two years. Over this period the computing resources of the EGEE Grid are expected to comprise approximately 9000 nodes and 1000Tb (Tb = 1024 gigabytes) of storage.

Why Bio-computing on the Grid?

Many tasks in Biology, medical image analysis or other health related applications are ideal candidates to make use of the specific qualities of Grid computing. The qualities of a Grid mean that it is inherently parallel (many simultaneous similar computations); it connects distributed resources and consumers; it provides large computing resources; it makes use of large amounts of data often held in heterogeneous, geographically-spread databases; and it allows scientists to collaborate in virtual communities.

Some examples of biological computing applications which require these characteristics include:

- Large data set analysis: for example, genomic analysis, population genetics, drug trials;
- Distributed applications: for example, accessing imaging technology for medical staff through portable computing, providing sophisticated computing “in the field”, interaction with distributed sequencing and proteomics facilities (creating workflows);
- Simulations: for example, protein folding applications, biochemical pathways and whole cell simulations, ecological simulations, biomedical visualisation (e.g. MRI simulations), ecological simulations;
- Heterogeneous data sets: for example, combining genomic, annotation and pathway information from existing bioinformatics databases.

EGEE and Biocomputing

Biomedical computing is one of the major areas identified by EGEE as an important community to help stimulate the uptake of Grid computing within the European Union (EU). This follows on from the experience of biomedical applications on the European DataGrid (EDG), a precursor project to EGEE. To facilitate the Biological community’s ability to participate in the project, a part of the EGEE project is dedicated to identifying and deploying biocomputing applications onto the EGEE Grid.

EGEE training

EGEE has a dedicated training activity with a remit to facilitate the entry into grid computing and its use by applications communities such as Biomedicine and Bioinformatics. All levels of courses can be provided, ranging from basic introductions to grid technology to expert application developer courses.

EGEE is a project funded by the European Commission - contract number INFSO-RI-508833

The training activity is led by the UK National eScience Centre in Edinburgh, with federated groups throughout Europe. To contact Training, email training-support@nesc.ac.uk or visit http://www.egee.nesc.ac.uk/ for further information.

What applications are available?
Applications in the Biomedical arena have been included in the EGEE project from the outset. The EGEE middleware will be built around standard Web Service technologies which are already widely understood and used in the biocomputing application development community.

- **GATE:** Radiotherapy planning, Medical tomography application, visit http://www-lphe.epfl.ch/~PET/research/gate/

- **GPS@:**
  Protein Grid sequence analysis – including blast, psi-blast, fasta, search, clustal, proscan, secondary structure prediction, SRS. Visit http://gpsa.ibcp.fr/

- **CDSS:**
  Clinical decision support. The CDSS aims at extracting medically relevant knowledge from a large set of information with the objective of guiding the practitioner in their clinical practice. A medical knowledge base is built from the data available on the Valencia regional hospital network. Visit http://egee-na4.ct.infn.it/biomed/20040608/NA4 Biomed Meeting - Pilots CDSS (UPV).ppt

- **PTM3D:**
  Interactive radiological image visualization and processing. PTM3D is addressing the problem of running interactive tasks on a grid for supervised medical data exploration and analysis.

- **Mammogrid:**
  The aim of this project, in light of emerging Grid technology, is to develop a European-wide database of mammograms that will be used to investigate a set of important healthcare applications as well as the potential of this Grid to support effective co-working between healthcare professionals throughout the EU. Visit http://mammogrid.vitamib.com/

EGEE is anxious to consider other applications. A form for candidate applications can be found at: http://egee-na4.ct.infn.it/biomed/process.html