

Application Architecture for High Performance Microarray Experiments over the Hellas-Grid Infrastructure

Wednesday, 9 May 2007 17:30 (20 minutes)

Describe the scientific/technical community and the scientific/technical activity using (planning to use) the EGEE infrastructure. A high-level description is needed (neither a detailed specialist report nor a list of references).

Microarray experiments permit a genome-scale evaluation of gene functions and are therefore among the most topical and prominent developments of biomedical research. Our target user group includes: Doctor, researchers, biologists, bioinformatics specialists carrying out DNA microarray experiments.

Report on the experience (or the proposed activity). It would be very important to mention key services which are essential for the success of your activity on the EGEE infrastructure.

Bioinformatics applications in general and microarray experiments in particular are perfectly tailored to Grid infrastructures. Motivated by this fact, University of the Aegean proposed to 'Gridify' a selected number of microarray analysis and normalization applications. These applications focus on cDNA arrays. The target Grid infrastructure was the Hellas Grid portion of the pan-European Grid infrastructure developed for e-science in the scope of the EGEE (Enabling Grids for E-Science in Europe) project and its successors. The idea was proposed in the scope of the Grid-App tender issued by the Greek General Secretariat of Research and Technology (GSRT) and was warmly embraced by the evaluators.

The application will leverage a host of Hellas Grid (i.e. gLite) middleware services including security, information services, data management, job submission, and resource management. The paper elaborates on the use of these services based on our experiences.

With a forward look to future evolution, discuss the issues you have encountered (or that you expect) in using the EGEE infrastructure. Wherever possi-

ble, point out the experience limitations (both in terms of existing services or missing functionality)

Our recently started developments involve the following steps:
I) Transforming our microarray applications (from MATLAB, which is not currently supported on HellasGrid.) to supported executables such as the open-source GNU Octave (forge) .
II) Exploiting appropriate gLite middleware components in the areas of security (GIS), information services (R-GMA), data management (LFC), job submission (WMS), and resource management.
III) Implementing an access portal based on UI images available.

Describe the added value of the Grid for the scientific/technical activity you (plan to) do on the Grid. This should include the scale of the activity and of the potential user community and the relevance for other scientific or business applications

Economic parameters are a major obstacle to microarray experiments, since some processes are extremely computationally demanding. This is particularly true for the normalization process, which deals with the fact that every repeat experiment will give rise to a certain amount of variations. Variations can be minimized based on statistical methods, which in turn allows one to compare the expression levels between multiple microarray experiments. Normalisation procedures rely on the fact that gene expression data can follow a normal distribution. Thus the entire distribution can be transformed about the population mean and median without affecting the standard deviation. cDNA arrays can contain up to 25,000 gene-complementary seqsequences while high-density oligonucleotide arrays can hold more than 100,000 proceses. Grid computing can accelerate the computations associated with microarray normalization, while at the same time providing access to vast amounts of federated storage.

Primary author: Prof. SOLDATOS, John (Athens Information Technology)

Co-authors: Dr CHATZIOANNOU, Aris (University of the Aegean); Prof. MAGLOGIANNIS, Ilias (University of the Aegean); Mr KANARIS, Ioannis (University of the Aegean); Mr MYLONAKIS, Vasileios (National Technical University of Athens)

Presenter: Prof. SOLDATOS, John (Athens Information Technology)

Session Classification: Poster and Demo Session