

Enabling Distributed Access and Parallel Processing of Bioinformatics Data in Grid Environment: The EKTORAS platform

Tuesday, February 12, 2008 4:00 PM (0 minutes)

Users, accessing the platform's Web interface through the implemented portal, are given the ability to submit their experiments, retrieve their results and also compare them with formerly submitted experiments. Access to services is enabled by parsing input files and accordingly activating the 'gridified' algorithms for processing the microarray experiments. Both data parsing operations and launching of experiments are specified as Grid jobs, using the Job Description Language (JDL). The provided microarray input files, which are usually structured according to formats that are standard for the microarray bioinformatics community, are pre-processed so as to be usable by the range of algorithms available. The results of this pre-processing step are directed to the Grid's storage elements (SE). Then the data are being processed by parallel applications distributing the parallel chunks & jobs to various nodes-processors of the Grid.

3. Impact

The described EKTORAS platform for microarray data analysis helps scientists and research groups with limited or no experience in microarray analysis to significantly reduce the processing time of large experiments. In the meantime the EKTORAS platform aims to bridge and integrate distributed databases storing multifaceted biological and medical information (gene sequences, homology across species, proteins encoded and relevant biochemical pathways) about gene function and structure, enabling ubiquitous access to bioinformatics data. The original architectural design will be constantly updated, based on the early feedback from experiments on the prototype infrastructure. However the already implemented modules have clearly proven the feasibility of the envisaged system.

URL for further information:

<http://www.icsd.aegean.gr/ektoras>, <http://hellasgrid.gr/>

4. Conclusions / Future plans

This work serves as a starting point for building a more complete and integrated Grid enabled microarray experimentation environment. In this context the EKTORAS access portal will be enhanced to allow end-users to retrieve experiment files from public biological databases such as the EBI microarray library (<http://www.ebi.ac.uk>). To this end the portal will provide an adapter to the EBI database system, allowing to view, browse and select EBI files and then process them in the Grid.

Provide a set of generic keywords that define your contribution (e.g. Data Management, Workflows, High Energy Physics)

cDNA microarray experiments, Bioinformatics, Data Management, MPI (Message Passing Interface)

1. Short overview

In this work we present a Web based platform called EKTORAS (<http://www.icsd.aegean.gr/ektoras>), which enables distributed access and parallel processing of biological data in Grid environments. The deployed software aims at creation of tools for processing data from microarray experiments over the Hellenic Grid infrastructure. This work serves as a starting point for building a more complete and integrated Grid enabled microarray experimentation environment.

Primary author: Prof. MAGLOGIANNIS, Ilias (University of Aegean)

Presenter: Prof. MAGLOGIANNIS, Ilias (University of Aegean)

Session Classification: Posters

Track Classification: Application Porting and Deployment