



Contribution ID: 64

Type: Demo

A bioinformatics platform: the evolution of the WISDOM environment

Tuesday, 23 September 2008 16:10 (0 minutes)

Describe the activity, tool or service using or enhancing the EGEE infrastructure or results. A high-level description is needed here (Neither a detailed specialist report nor a list of references is required).

Mature and efficient, the WISDOM environment has evolved continuously to take advantage of the Grid computing and storage resources for large-scale drug discovery applications. The environment has also been used on several occasions for other bioinformatics applications, thanks to its flexibility and its simplicity. This new bioinformatics platform just brings the environment to a new level: combining bioinformatics tools altogether within the same framework.

Report on the impact of the activity, tool or service. This should include a description of how grid technology enabled or enhanced the result, or how you have enabled or enhanced the infrastructure for other users.

Based on the experience gained with the WISDOM environment, the platform makes an efficient use of the grid services. Thanks to the AMGA metadata catalogue it uses a global information system to share all the biological data used by its services for better interoperability. The system is based on the use of agents, or software wrappers; the agents are designed to run any software integrated in the platform and pick up tasks on the fly. All the information concerning the tasks is stored in the platform information system, especially information on input data needed by the jobs and some important results as well as the links to the outputs stored on the grid. The jobs also save some statistics dynamically in AMGA. The automatic replication and update service is based on the grid data management system, using LFC and storage elements to manage the physical files and manage the metadata in AMGA.

Abstracts for online demonstrations must provide a summary of the demo content. Places for demos are limited and this summary will be used as part of the selection procedure. Please include the visual impact of the demo and highlight any specific requirements (e.g. network connection). In general, a successful demo is expected to have some supporting material (poster) and be capable of running on a single screen or projector.

In this demonstration, we will show how the different tools can be simply managed through external web pages for simple invocation, or in a more complex way through workflow engines (MOTEUR, Taverna) while monitoring in real time what is actually happening on the grid. The demonstration will focus on user-friendliness and versatility of the platform as its tools can be used transparently and efficiently. It will show on a single screen some use cases, and show how the platform can be used for both simple short deadline jobs and large-scale deployments. Also we will demonstrate the interoperability of the platform tools with

external third party clients like Taverna or Moteur. For this demonstration we foresee to use a screen and internet connection to access the services of the platform.

Describe the added value of the grid for your activity, or the value your tool or service adds for other grid users. This should include the scale of the activity and of the potential user community, and the relevance for other scientific or business applications.

The platform has been developed with one idea in mind: addressing the needs of the bioinformatics community that work with workflows and that may need the computing power of the grid or just need computers to run their tasks. By the use of agents, the platform allows quick submission of jobs and makes the grid useful and efficient for short jobs and large-scale deployments alike. The use of an interoperable technology also enhances the integration capabilities of new software and databases, that can easily be used altogether through workflow engines. Advanced tools used in the platform manage automatically the job submissions, replication and update of the biological databases so end-users just have to focus on their workflows and not on the technology. Through its use of Web Services interfaces, which is backed by the European bioinformatics community, the platform will provide new solutions for accessing the grid to people that are not familiar with it.

Primary authors: Mr SALZEMANN, Jean (CNRS/IN2P3); Mr REICHSTADT, Matthieu (CNRS/IN2P3); Mr BLOCH, Vincent (CNRS/IN2P3)

Co-authors: Ms DA COSTA, Ana Lucia (CNRS/IN2P3); Ms DELAHAYE, Manon (CNRS/IN2P3); Mr TRAN, Tuan Tu (CNRS/IN2P3); Mr BRETON, Vincent (CNRS/IN2P3)

Presenters: Mr SALZEMANN, Jean (CNRS/IN2P3); Mr REICHSTADT, Matthieu (CNRS/IN2P3); Mr BLOCH, Vincent (CNRS/IN2P3)

Session Classification: Demos and Posters

Track Classification: Demo