



Contribution ID: 48

Type: Poster

G-PPI: Discovery of Protein-To-Protein Interactions On the Grid

Tuesday, 23 September 2008 16:24 (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).

We have gridified an existing parallel application for discovering protein-to-protein interactions in a protein database using the P-GRADE portal as part of the SEE-GRID 2 project.

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.

The common grid applications seem to be MPI applications encapsulated in a JDL file. However, we think that decomposing the parallel task to several small non-MPI jobs can be more effective and suitable for the Grid. Our application demonstrates that this is indeed possible, using the high level GUI of P-GRADE. We think that our application should encourage better infrastructure for supporting a large number of jobs and fault recovery, and for designing applications in this vein. In particular, we have shown how useful P-GRADE can be for application developers.

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.

We have made use of the Parameter Study support in P-GRADE, and we have shown how to gridify a real-world application using the high-level interface of P-GRADE alone, by implementing a map/reduce style computation using the Generator/Collector jobs in P-GRADE's PS feature. The simplicity of the application and the way we have managed to encapsulate the gridification in the P-GRADE workflow should be of appeal to grid application developers. The aim of the gridification is to speed up the massive computations required by large protein databases by distributing the jobs to grid sites, making use of multiple sites for the computation.

Primary authors: GURSOY, Attila (Koc University); ULUBAS, Cengiz (Koc University); OZKURAL, Eray (Bilkent University)

Presenter: OZKURAL, Eray (Bilkent University)

Session Classification: Demos and Posters

Track Classification: Poster