

Experiences from porting the astrophysical simulation “The unified theory of Kuiper-belt and Oort-cloud formation” to EGEE grid

Tuesday, 12 February 2008 12:00 (20 minutes)

The experiment was performed in the scope of collaboration between Astronomical Institute of Slovak Academy of Sciences, Catania Observatory and Adam Mickiewicz University in Poznan. The simulation was ported to EGEE by Institute of Informatics Slovak Academy of Sciences and it ran in EGEE and TriGrid from February to October 2007.

The simulation consists of a sequence of sub-simulations with many independent tasks within each sub-simulation. The necessary requirement is to finish all the tasks of a given sub-simulation before starting the next sub-simulation.

The main problem when running the large number of jobs in grid was the reliability of grid infrastructure. Job management was rather time consuming due to the time spent on the analysis of the failed jobs and their resubmission. Moreover, the jobs that were waiting at some sites in a queue for a long time were blocking whole simulation.

3. Impact

To overcome these problems we developed an easy-to-use framework based on “pilot jobs” concept that uses only services and technologies available in EGEE. It consists of pilot jobs (“workers”) and automatic job management script.

Workers are running the application code in cycle with input datasets downloaded from Storage Element using RFIO access. Output datasets are stored in output folder. To check the progress, the user only needs to list the contents of the output folder. To identify hanging jobs or the jobs that performs too slowly, the workers are periodically sending a monitoring information to SE (“heart beat”). To avoid termination of workers by queuing system, the workers are running only for limited time.

The main goal of the job management script is to maintain the defined number of active workers with detection of failed submissions, finished and waiting workers. It uses job collections to speedup the startup and automatic blacklisting of full and erroneous sites.

4. Conclusions / Future plans

One of the expectations of grid users is that they just put their application code and input data into the grid, configure and start the processing and after the processing (with occasional checking the progress) they download the output data. In our approach we tried to get as close as possible to this expectation. The users of the astrophysical application were satisfied with our framework and we plan to use it for porting of similar applications to EGEE.

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

astrophysical simulation, parameter study

1. Short overview

Main goal of the simulation was to work out a unified theory of the formation of all: Jovian planets, Kuiper belt, Scattered Disc (populations of small bodies beyond the Neptune’s orbit) and Oort cloud. The simulation was based on the dynamical evolution of a large number (~10000) planetesimals treated as test particles in the proto-planetary disc. The main reason for using the grid was the need for about 40 CPU-years of computing time.

Primary author: Mr ASTALOS, Jan (Institute of Informatics, Slovak Academy of Sciences)

Co-authors: Dr HLUCHY, Ladislav (Institute of Informatics, Slovak Academy of Sciences); Mr DOBRUCKY, Miroslav (Institute of Informatics, Slovak Academy of Sciences)

Presenter: Mr ASTALOS, Jan (Institute of Informatics, Slovak Academy of Sciences)

Session Classification: Astronomy & Astrophysics

Track Classification: Application Porting and Deployment