Contribution ID: **65**   Type: **Oral Presentation**

Theompi: a large MPI cluster on the grid for Theoretical Physics

Wednesday, 13 April 2011 16:30 (30 minutes)

Overview

Theompi is a project aimed at the deployment of a large MPI cluster on the grid, available for the whole INFN theoretical physics community.

The cluster has been installed at the INFN Pisa site and is Grid-enabled though a CREAM based computing element. The production version of the software has been enhanced with an experimental patch to better support and customize the parallel job execution on a multicore environment, according to the recommendations given by the EGEE MPI WG.

This patch has been tested and deployed in collaboration with the gLite middleware developers.

We describe the porting of significant applications in Theoretical Physics, executed in this new environment and taking advantage of the new grid parallel attributes.

Impact

The INFN theoretical physics community consists of 700 FTE researchers distributed over 28 sites and involved in 60 research projects. A few projects, molstly concerning lattice simulations (LQCD, fluid dynamics, ..), require massive HPC systems with next neighbors high speed inter-connections. Most other projects make deep usage of standard HPC resources, often provided in the past by small or medium sized local clusters. The Theompi project has provided the community with a transparent and flexible mechanism to share HPC resources.

Description of the work

In recent years the Egee/Egi grid infrastructure has proven to be a robust and scalable infrastructure for sequential scientific computation, but hardly usable for parallel programming.

In June 2010 the EGEE MPI WG has finalized a proposal aiming at a more flexible slot allocation for parallel jobs. In particular the support for the JDL attributes (to specify in detail how the slots needed for the execution of a parallel application should be allocated) was proposed.

A preliminary patch supporting these new attributes has been provided by gLite middleware developers, enabling the possibility of a real operational usage of the grid infrastructure for parallel applications.

This recent evolution has convinced the INFN theoretical physics community to install a first large grid-enabled MPI cluster.

The installation, deployed at the INFN Pisa site, is composed by 1024 Opteron cores and Infiniband high speed communication

infrastructure. The hardware is in place since summer 2010 and

the MPI support is operational since december 2010.

URL

Conclusions

The support of the new granularity attributes in the gLite middleware gives a real chance to use grid clusters for both MPI and multithreaded applications. We decided to start with a single large cluster, but the flexibility of this mechanism will allow the integration of other (old and new) clusters in a near future.

Primary author: ALFIERI, Roberto (INFN Parma - Parma University)

Co-authors: CIAMPA, Alberto (INFN Pisa); GIANELLE, Alessio (INFN Padova); MAZZONI, Enrico (INFN Pisa); DI RENZO, Francesco (INFN Parma - Parma University); ANDRONICO, Giuseppe (INFN Catania); SGARA-VATTO, Massimo (INFN Padova); DE PIETRI, Roberto (INFN Parma - Parma University); MONFORTE, Salvatore (INFN Catania); AREZZINI, Silvia (INFN Pisa)

Presenter: ALFIERI, Roberto (INFN Parma - Parma University)

Session Classification: User Environments

Track Classification: High Performance Computing - MPI Environments