Contribution ID: **28** Type: **Oral Presentation**

A uniform MPI user experience with MPI-Start in EMI

Monday, 11 April 2011 12:00 (30 minutes)

Overview

Execution of parallel applications on a grid environment is a key task which requires the cooperation of several middleware tools and services. Although current middleware stacks provide basic access to such applications, this support results to be limited in practice. EMI wants to face the challenge of providing extended and seamless access to the execution of MPI jobs across the middleware it supports. This is being done building on an existing project, MPI-Start, originally implemented in the context of the int.eu.grid and which has been used in the gLite/EGEE Grid over the last years. In this work we present the effort EMI is putting on MPI-Start, that is targeted on the provision of a uniform abstraction for the execution of parallel jobs and the integration of the tool with all the middleware stacks.

Impact

Users communities have frequently stated that the support for MPI applications is too coarse and unwieldy for their needs. The processes of job submission and management should be simple, more transparent and should support generic parallel job types.

The adaptation and integration MPI-Start in the middleware stacks of EMI was performed by using the particular mechanisms of each middleware - Runtime Environments of ARC, Execution Environments of UNICORE and user wrapper scripts in gLite. Therefore users get a uniform experience for running their jobs with MPI-Start, while at the same time they maintain the interface for submitting the jobs of their middleware.

In most cases the automatic environment detection mechanisms of MPI-Start provide support for the most common MPI implementations and resource managers without any administrator intervention, thus allowing sites with little or no previous MPI experience to support these applications without major efforts.

The possibility of defining how user processes are mapped to physical resources and the modular architecture of MPI-Start, allows the execution of new kind of parallel jobs. OpenMP jobs are now supported using the common interface of MPI-Start.

Description of the work

Middleware support for MPI applications is usually limited to the possibility of allocating a set of nodes. The user still needs to deal with low level details that make the task non trivial. Furthermore, the heterogeneity of resources available in Grid infrastructures aggravates the complexity that users must face to run their applications.

MPI-Start provides users with an abstraction layer that simplifies the execution of MPI and other types of parallel applications on Grid resources. By using a modular and pluggable architecture it manages for the user the details of:

* Local Resource Management Systems
* File Distribution
* Application Compilation
* Application Execution

Although MPI-Start is designed to be middleware agnostic, its original target was the gLite environment. In this work we have integrated it into ARC and UNICORE middlewares using their specific mechanisms, Runtime Environments of the ARC CE and Execution Environments of UNICORE.

The latest developments of MPI-Start have also introduced a new architecture for extensions that allows further customize the way the site and user are able to modify the MPI-Start behavior.

Moreover, MPI-Start now supports basic mapping of logical processes to physical resources by allowing the user to specify how many processes want to execute for each of the available hosts for their execution.

Conclusions

The unique interface of MPI-Start simplifies the task of starting the jobs by handling transparently the low level details of the different Local Resource Management Systems, execution frameworks and file distribution methods. The most common batch systems and MPI implementations are supported by MPI-Start and its modular architecture allows the easy extension of the tool to support new implementations.

With the integration of MPI-Start into the three computing middleware stacks of EMI, users have a unified experience for running their parallel jobs.

The availability of multi-core architectures opens the possibilities of new types of parallel jobs. MPI-Start ability to define the way the user logical processes introduces the possibility of new types of applications in the tool, such as hybrid MPI/OpenMP applications.

Primary author: Dr FERNANDEZ DEL CASTILLO, Enol (CSIC)

Presenter: Dr FERNANDEZ DEL CASTILLO, Enol (CSIC)

Session Classification: EMI: Software for Distributed Computing Infrastructures

Track Classification: High Performance Computing - MPI Environments