Contribution ID: **116** Type: **Oral**

Distributed Task Scheduling for Physics Fusion Applications

Monday, February 11, 2008 4:20 PM (20 minutes)

There are two kinds of parallel loop schedulers to distribute program loops among the processors of a parallel architecture: static and dynamic scheduling. In this work, we will focus on dynamic schedulers because they are more suitable for heterogeneous environments such as a Grid. In general, in these algorithms a central node dynamically distributes a fraction of the computational workload (chunk) to the rest of the worker nodes. Depending on how the chunks are calculated, different simple self-scheduling schemes can be devised. An alternative to these schemes is the distributed self-schedulers. In this case, the scheduler takes into account the characteristics of the different components of the system (e.g. the cpu speed or the system load) to optimize the chunk assigned to each node. This work presents a new distributed self-scheduler scheme that takes into account all Grid characteristics: a high degree of heterogeneity, high fault rate, dynamic resource availability, etc.

3. Impact

The effects of this new distributed algorithm will be proved in the MARATRA (MAssive RAy TRAcing in Fusion Plasmas) system. MARATRA aids those community members who are working on the optimization of plasma heating by electron Bernstein waves (EBW). This new algorithm allows the execution of tasks in the MARATRA system using loop parallelization methods. This approach presents important advantages over the traditional task schedulers, for example, a better workload balancing between all Grid resources or a decrease of the scheduling overhead. Furthermore, the estimated execution time of each Grid node during the tasks distribution process allows the dynamically adaptation of the whole application. Hence, the workload of each task will be dynamically distributed depending on the behaviour of each node. The goal of this distribution scheme is to adapt the MARATRA system to the Grid environment.

4. Conclusions / Future plans

The high degree of heterogeneity and high fault rate of existing grid infrastructures require the implementation of new self-scheduling algorithms to calculate the task chunk size of parameter sweep and high throughput computing applications. The presentation will describe a new algorithm inspired in the distributed self-schedulers schemes used for loop distribution on parallel architectures. Its advantages are demonstrated for the execution of a Physics Fusion application.

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

SelfSchedulers, Grid, Dynamic, Distributed, Fusion, Algorithm

1. Short overview

A new scheduling algorithm to distribute tasks on Grid environments will be described. The algorithm is an enhancement of the distributed dynamic self-scheduler algorithm used in loop parallelization. The algorithm will be applied to the efficient distribution of tasks in physics fusion simulation codes.

Primary author: Mr HERRERA, Jose (Teaching Assistant)

Co-authors: Dr HUEDO, Eduardo (Associate Professor); Prof. LLORENTE, Ignacio Martín (Full Professor); Dr MONTERO, Ruben Santiago (Associate Professor)

Presenter: Mr HERRERA, Jose (Teaching Assistant)

Session Classification: Fusion

Track Classification: Existing or Prospective Grid Services