



Predicting Resource Requirements of a Job Submission

Wednesday 29 September 2004 10:00 (1 minute)

Grid computing provides key infrastructure for distributed problem solving in dynamic virtual organizations. However, Grids are still the domain of a few highly trained programmers with expertise in networking, high-performance computing, and operating systems.

One of the big issues in the full-scale usage of a grid is the matching of the resource requirements of a job submission to available resources. In order for resource brokers/job schedulers to ensure efficient use of grid resources, an initial estimate of the likely resource usage of a submission must be made. In the context of the Grid Enabled Analysis Environment (GAE), physicists want the ability to discover, acquire, and reliably manage computational resources dynamically, in the course of their everyday activities. They do not want to be bothered with the location of these resources, the mechanisms that are required to use them, keeping track of the status of computational tasks operating on these resources, or with reacting to failure. They do care about how long their tasks are likely to run and how much these tasks will cost.

So the grid scheduler must have the capability to estimate before job submission, how much time and resources the job will consume on execution site. Our proposed module, Prediction engine will be part of scheduler and it will provide estimates of resource use along with the duration of use. This will enable scheduler to choose the optimum site for job execution.

This paper presents the survey of existing grid schedulers and then based on this survey states the need for resource usage estimation. Also the architecture and design of "grid prediction engine" that predicts the resource requirements of a job submission is discussed.

Authors: ALI, A. (NIIT); ANJUM, A. (NIIT); MEHMOOD, A. (NIIT); NEWMAN, H. (Caltech); WILLERS, I. (Caltech); BUNN, J. (Caltech); MECLATCHEY, R. (UWE)

Presenter: ANJUM, A. (NIIT)

Session Classification: Poster Session 2

Track Classification: Track 4 - Distributed Computing Services