Contribution ID: **48** Type: **Oral**

Execution Time Prediction of Imperative Paradigm Tasks for Grid Scheduling Optimization

Wednesday 13 February 2008 14:20 (20 minutes)

The proposed system is intended to be implemented in the KnowledgeGRID Malaysia to improve the efficiency of the scheduling system. The work is focused on imperative paradigm tasks since they are commonly used in the aforementioned grid. Imperative paradigm refers to a sequence of commands for the computer to perform and the normally used imperative paradigm programming languages are R!, Fortran and C. In the current phase, all testing and evaluation is done via a web-based wrapper which is developed specially for this purpose. The testing and evaluation involves test cases sampled from jobs submitted to the aforesaid grid.

3. Impact

We proposed a novel methodology and architecture to predict the execution time of jobs. In this phase, the proposed prediction module works as a standalone system which would estimates the execution time of jobs to assist scheduling in existing middleware used in the grid. A mathematical model and benchmarked data are used to forecast the time required to execute jobs. An incoming job is categorized according to its application, and is then parsed and broken down into smaller units known as tokens. The complexity and relationship amongst these tokens are then analyzed. The execution times for the tokens are then combined to give an estimate of the execution time of the entire job. An accurate estimation of jobs'execution time in advance allows allocation of resources into appropriate queues, which eventually leads to effective scheduling. Also, a mathematical model has been developed for the purpose of comparing the theoretical optimized scheduling system with that of the actual grid.

4. Conclusions / Future plans

The experimental results from the sampled test cases and developed prototype show that the technique is successful in achieving an accuracy of greater than 80%. As this work focuses on imperative paradigm programming, perhaps future work may suitably involve other paradigms such as object-oriented and data-intensive programming. Also, further research may look into integrating the prediction module into the real grid environment instead of a standalone, web-based system.

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

Prediction module, Grid scheduling, Job execution time, Grid queue length distribution

1. Short overview

An efficient functioning of a complicated grid environment requires a resource manager to monitor the idling resources and to schedule users' submitted jobs accordingly. However, at present, the execution time prediction depends mostly on pure guesswork. The inaccuracy of guesswork leads to inefficient resource usage, incurring extra operation costs. Thus, we propose a job execution time prediction module that estimates the execution time of jobs to optimize the scheduling system in the grid.

Authors: Ms LIM, Mei Kuan (Mimos Berhad); Dr YAP, Yee Jiun (Mimos Berhad)

Co-authors: Ms KIRAN, Maleeha (Mimos Berhad); Dr LAI, Weng Kin (Mimos Berhad)

Presenters: Ms LIM, Mei Kuan (Mimos Berhad); Dr YAP, Yee Jiun (Mimos Berhad)

Session Classification: Workflow and Parallelism

Track Classification: Existing or Prospective Grid Services