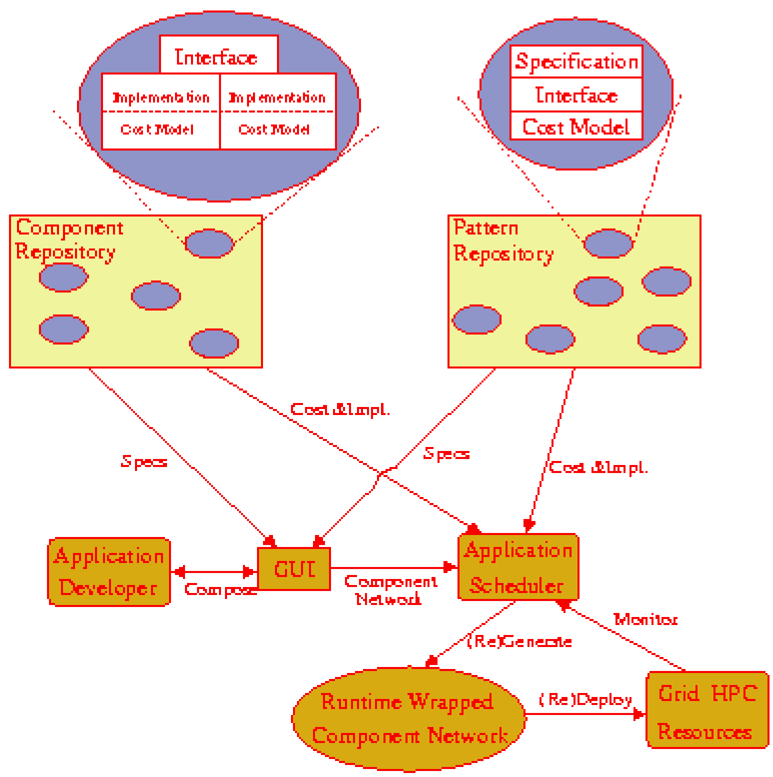
Enhance

Enhancing the Performance Predictability of Grid Applications
with Patterns and Process Algebras
tuyt

Overview

One of the most promising technical innovations in present-day computing is the invention of Grid technologies which harness the computational and storage power of widely distributed collections of computers. Grid technologies are breaking new ground in e-Science where scientists across the globe can collaborate and share data sets, processing power and specialised scientific instruments to accelerate the pace of scientific development. Grid technologies are also opening new doors in e-Business where small or medium-scale businesses can now have access to super-compting power which was once the province of only the most wealthy banks and multi-nationals. Data and resource sharing schemes such as these are changing the basic ground rules of computing.



All such schemes raise difficult issues of resource allocation and scheduling (roughly speaking, how to decide which computer does what, and when, and how they interact) which are made all the more complex by the inherent unpredictability of resource availability and performance. For example, I may forget to switch on my PC, a supercomputer may be required for a more important task, the internet connections I need may be particularly busy.

The Enhance project aims to simplify the effective programming of such systems by exploiting and synthesizing results from two underlying research programmes. Stochastic process algebras such as PEPA are used to model the behaviour of concurrent systems in which some aspects of behaviour are not precisely predictable. Pattern based programming recognizes that many real applications draw from a range of well known solution paradigms and seeks to make it easy for an application developer to tailor such a paradigm to a specific problem without re-inventing the wheel.

The choice of a particular paradigm carries with it considerable information about implied scheduling dependencies. By modelling these with PEPA, and thereby being able to include aspects of uncertainty which are inherent to Grid computing, we believe that we will be able to underpin systems which can make better scheduling and dynamic rescheduling decisions than less sophisticated approaches.

About the project

The vision of Grids as both tools and enabling fabrics for distributed collaborative computation continues to excite applications designers and challenge the Computer Science research community. While some of the issues raised are familiar from previous environments in distributed and parallel computing, many are new.

This project seeks to address the allied problems of programmability, cost predictability and scheduling by augmenting a component based Grid HPC framework with a mechanism for constructing applications as instantiations of pre-defined application patterns. Each pattern will encapsulate the logical behaviour of a familiar pattern of multi-component interaction, at a level of abstraction which simplifies the developer's task, without over constraining the selected implementation.

The definition and manipulation of semantic and performance behaviour of components and patterns will be facilitated by the use of a performance oriented process algebra as a modelling tool. Underpinning an application development process of the form illustrated below, our key insights are:

•that the use of patterns will helpfully constrain the implementation challenge, by providing good knowledge of the overall evolution of an application's interactions;

•correspondingly, that patterns allow both more accurate cost prediction (which in turn informs good scheduling decisions) and lighter weight monitoring (because the pattern based model tells us what the crucial aspects are for each instantiation);

•that cost modelling with a formalised, stochastic performance process algebra and associated tools facilitates capture and manipulation of the remaining uncertain aspects of application behaviour.

Ultimately, by construction of a prototype and experimentation with a number of case studies, we will demonstrate that this approach allows simple, intuitive construction of applications, and facilitates run-time scheduling and rescheduling with an effectiveness which surpasses that of less constrained methodologies.

Personnel

Anne Benoit, Murray Cole, Stephen Gilmore, Jane Hillston

Funding

The Enhance project is funded by the Engineering and Physical Sciences Research council grant number GR/S21717/01. The project commences on June 1st 2003 and will end on May 31st 2006.