HETEROGENEOUS PARALLEL-DISTRIBUTED PROCESSING APPLIED TO PROCESS ENGINEERING

Ph. D. Thesis in Computer Science

Vazquez, Gustavo Esteban
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
Avda. Alem 1253, 8000 Bahía Blanca, ARGENTINA
Email: gvazquez@criba.edu.ar

Supervised by:

Brignole, Nélida Beatriz
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
Avda. Alem 1253, 8000 Bahía Blanca, ARGENTINA
Planta Piloto de Ingeniería Química (UNS-CONICET)
Km 7 Cno. La Carrindanga, 8000 Bahía Blanca, ARGENTINA
Email: dybrigno@criba.edu.ar

First presented in:

July, 2002

THESIS OVERVIEW

The main goal of this thesis was to design new parallel processing strategies specially conceived for distributed environments in order to solve numerical and structural problems from the field of process systems engineering more efficiently. More specifically, the numerical sample problem addressed in this work was the optimization of nonlinear objective functions subjected to sets of nonlinear constraints, while the structural sample problem was the development of parallel-distributed structural techniques for process instrumentation design. These two kinds of problems are completely different as regards algorithmic nature as well as information types and handling. At the same time, the scope is ample because the basic theory that support these developments is of widespread use. Therefore, the design methodologies and performance analyses presented in this thesis are also illustrative of how to deal with other problems with similar features that may arise in various fields well beyond process systems engineering.

Broadly speaking, two general research lines for the development of parallel-distributed algorithms were addressed: the parallelization of existing serial algorithms and the creation of inherently parallel alternatives. For structural problems, the classic sequential methods for depth-first searches along undirected graphs were studied and some limitations for their use in networks of workstations were individualized. On the basis of this analysis, a new semi-dynamic load-balancing method was proposed and applied to the GS-FLCN algorithm for observability analysis. As regards inherently parallel algorithms, a novel totally distributed search algorithm that achieves greater efficiency in path exploration was devised for the same specific application.

With respect to the numerical problems, we designed and implemented various strategies to apply parallelism to the computationally intensive sections of the two most successful existing sequential algorithms for nonlinear constrained optimisation, namely GRG (Generalised Reduced Gradient) and SQP (Successive Quadratic Programming). Another original contribution in this area is the development of an optimisation technique that employs domain decompositions. The method is based on an inherently parallel algorithm that had been originally conceived for unconstrained problems. In particular, two effective strategies for the distribution of the variables among the blocks could be defined. This proposal succeeds in expanding the applicability range of the basic technique so that it can be employed for the efficient treatment of many complex nonlinear constrained optimisation problems that often arise in process engineering.

As to the nature of the processors, both homogeneous and heterogeneous clusters were considered. In general, homogeneous runs were carried out first for the sake of simplicity. In some cases, scale-up studies were performed using up to ten processors. For the heterogeneous trials, the clusters comprised three different kinds of processors. It is interesting to remark that the adoption of a semi-dynamic load-balancing approach proved to accommodate very well to heterogeneous environments. For the performance analysis, the classic speed-up metrics had to be adapted in order to ensure fair comparisons that took into account the heterogeneity of the processors. In this sense, all the proposed algorithms exhibited satisfactory results in comparison with the corresponding sequential run times.

Algorithmic performance was tested on academic and industrial sample problems. Several case studies were mathematical models of real items of engineering equipment or representations of existing industrial process plants. Due to the size and complexity of the rigorous mathematical models that were associated to industrial plants, it was necessary to develop a model generator (ModGen) that included special features to support instrumentation design packages. This a CAD tool that provides user-friendly facilities for the fast definition of plant topology and the automatic translation of the flowsheet information into the desired mathematical model.

In short, this thesis provides insight into the use of clusters for parallel processing applications to practical problems whose solution relies on methods based on either graph theory or nonlinear optimisation. In this context, useful new alternatives especially conceived for distributed-memory environments were proposed and successfully evaluated.