Chapter 17

Communication Issues in Scalable Parallel Computing

C. E. R. Alves
Universidade Sao Judas Tadeu, Brazil

E. N. Cáceres
Universidade Federal de Mato Grosso do Sul, Brazil

F. Dehne
Carleton University, Canada

S. W. Song
Universidade de Sao Paulo, Brazil

ABSTRACT

In this book chapter, the authors discuss some important communication issues to obtain a highly scalable computing system. They consider the CGM (Coarse-Grained Multicomputer) model, a realistic computing model to obtain scalable parallel algorithms. The communication cost is modeled by the number of communication rounds and the objective is to design algorithms that require the minimum number of communication rounds. They discuss some important issues and make considerations of practical importance, based on our previous experience in the design and implementation of parallel algorithms. The first issue is the amount of data transmitted in a communication round. For a practical implementation to be successful they should attempt to minimize this amount, even when it is already within the limit allowed by the CGM model. The second issue concerns the trade-off between the number of communication rounds which the CGM attempts to minimize and the overall communication time taken in the communication rounds. Sometimes a larger number of communication rounds may actually reduce the total amount of data transmitted in the communications rounds. These two issues have guided us to present efficient parallel algorithms for the string similarity problem, used as an illustration.

DOI: 10.4018/978-1-60566-661-7.ch017
Communication Issues in Scalable Parallel Computing

INTRODUCTION

In this book chapter, we discuss some important communication issues to obtain a highly scalable computing system. Scalability is a desirable property of a system, a network, or a process, which indicates its ability to either handle growing amounts of work in a graceful manner, or to be readily enlarged. We consider the CGM (Coarse-Grained Multicomputer) model, a realistic computing model to obtain scalable parallel algorithms. A CGM algorithm that solves a problem of size \( n \) with \( p \) processors each with \( O(n/p) \) memory consists of an alternating sequence of computation rounds and communication rounds. In one communication round, we allow the exchange of \( O(n/p) \) data among the processors. The communication cost is modeled by the number of communication rounds and the objective is to design algorithms that require the minimum number of communication rounds. We discuss some important issues and make considerations of practical importance, based on our previous experience in the design and implementation of several parallel algorithms.

The first issue is the amount of data transmitted in a communication round. For a practical implementation to be successful we should attempt to minimize this amount, even when it is already within the maximum allowed by the CGM model which is \( O(n/p) \).

The second issue concerns the trade-off between the number of communication rounds which the CGM attempts to minimize and the overall communication time taken in the communication rounds. Under the CGM model we want to minimize the number of communication rounds so that we do not have to care about the particular interconnection network. In a practical implementation, we do have more information concerning the hardware utilized and the communication times in a particular interconnection network. Sometimes a larger number of communication rounds may actually reduce the total amount of data transmitted in the communications rounds. Although the goal of the CGM model is to minimize the number of communication rounds, ultimately the main objective is to minimize the overall running time that includes the computation and the communication times.

These two issues have guided us to present efficient parallel algorithms for the string similarity problem, used as an illustration. By using the wavefront-based algorithms we present in this book chapter to illustrate these two issues, we also address a third issue, the desirability of avoiding costly global communication such as broadcast and all-to-all primitives. This is obtained by using wavefront or systolic parallel algorithms where each processor communicates with only a few other processors.

The string similarity problem is presented here as an illustration. This problem is interesting in its own right. Together with many other important string processing problems (Alves et al., 2006), string similarity is a fundamental problem in Computational Biology that appears in more complex problems (Setubal & Meidanis, 1997), such as the search of similarities between bio-sequences (Needleman & Wunsch, 1970; Sellers, 1980; Smith & Waterman, 1981). We show two wavefront parallel algorithms to solve the string similarity problem. We implement both the basic algorithm (Alves et al., 2002) and the improved algorithm (Alves et al., 2003) by taking into consideration the communication issues discussed in this book chapter and obtain very efficient and scalable solutions.

PARALLEL COMPUTATION MODEL

Valiant (1990) introduced a simple coarse grained parallel computing model, called Bulk Synchronous Parallel Model – BSP. It gives reasonable predictions on the performance of the algorithms when
Related Content

Analyzing the Cooling Efficiency of the Thermal Environment in Data Centers

Efficient Communication Interfaces for Distributed Energy Resources
[www.igi-global.com/chapter/efficient-communication-interfaces-distributed-energy/61991?camid=4v1a](www.igi-global.com/chapter/efficient-communication-interfaces-distributed-energy/61991?camid=4v1a)

Towards Energy Efficiency for Cloud Computing Services
[www.igi-global.com/chapter/towards-energy-efficiency-for-cloud-computing-services/82544?camid=4v1a](www.igi-global.com/chapter/towards-energy-efficiency-for-cloud-computing-services/82544?camid=4v1a)

An Analytical Model for Resource Characterization and Parameter Estimation for DAG-Based Jobs for Homogeneous Systems