

Applications and Development of Communication Models
for the Touchstone GAMMA and DELTA Prototypes

FINAL REPORT

covering the period
January 1, 1992 - June 30, 1993

Steven R. Seidel
Principal Investigator

Department of Computer Science
Michigan Technological University
1400 Townsend Drive
Houghton, Michigan 49931-1295

NASA Ames Research Center Grant Number NAG 2-757

(Submitted August 16, 1993)

(NASA-CR-193617) APPLICATIONS AND
DEVELOPMENT OF COMMUNICATION MODELS
FOR THE TOUCHSTONE GAMMA AND DELTA
PROTOTYPES Final Report, 1 Jan.
1992 - 30 Jun. 1993 (Michigan
Technological Univ.) 4 p

N94-16867

Unclas

G3/32 0179689

1 Project Summary

The goal of this project was to develop models of the interconnection networks of the Intel iPSC/860 and DELTA multicomputers to guide the design of efficient algorithms for interprocessor communication in problems that commonly occur in CFD codes and other applications. Interprocessor communication costs of codes for message-passing architectures such as the iPSC/860 and DELTA significantly affect the level of performance that can be obtained from those machines. This project addressed several specific problems in the achievement of efficient communication on the Intel iPSC/860 hypercube and DELTA mesh. In particular, an efficient global processor synchronization algorithm was developed for the iPSC/860 and numerous broadcast algorithms were designed for the DELTA. This work is described in more detail below.

One goal of this project was to improve communication performance in areas identified by experience with the development of CFD codes for the iPSC/860 hypercube. The basic communication problem of interest in this case was the “shift” operation, where each processor sends a message to its neighbor in a ring. A detailed communication model was developed in [1, 3, 4, 5] to show why the shift operation was slower than predicted by earlier communication models. It was demonstrated that globally synchronizing the processors was necessary to achieve the most efficient performance of the shift operation. A major part of this work was the development of a global processor synchronization algorithm that synchronized the processors more precisely than other currently available algorithms.

Another goal of this project was to develop a communication model for the DELTA mesh. This work started with a study of the model developed by Robert van de Geijn, Rik Littlefield, and others. The broadcasting problem was chosen as the vehicle used to validate that model. Many variations of some basic broadcasting algorithms were developed and tested. Some of these algorithms performed better than those given earlier by van de Geijn. In most cases it was found that ordinary programming practices were not sufficient to achieve the best communication performance from the DELTA and that the current version of the communication model does not adequately explain why this is the case. This work is reported in [2]. Further refinement of the communication model of the DELTA is required to accurately predict the costs of communication operations. However, in light of the recent

availability of the Intel PARAGON, this work should be continued on that machine.

The bibliography at the end of this report shows where this work has appeared in the general literature. Item [2] is currently being prepared for journal submission. Following are abstracts of the two major publications that have resulted from this work so far [2, 3].

2 Abstracts

2.1 Global synchronization algorithms for the Intel iPSC/860

In a distributed memory multicomputer that has no global clock, global processor synchronization can only be achieved through software. Global synchronization is used in many applications, including tridiagonal systems solvers, CFD codes, and sequence comparison algorithms. For the Intel iPSC/860 in particular, global synchronization can also be used to ensure the most effective use of the communication network. Three global synchronization algorithms are considered for the iPSC/860: the **gsync** primitive provided by Intel, the PICL primitive **sync0**, and the RDS algorithm. Based on the communication model presented here, it is shown that **gsync** sometimes leaves the processors more poorly synchronized than they were to begin with. It is also shown that interrupts from the node operating system can cause **gsync** to contend for communication ports with the application code. The RDS algorithm does not have these shortcomings and costs only slightly more than the other algorithms. Measurements of the cost of message shift operations preceded by global synchronization confirm that the RDS algorithm always synchronizes the nodes more precisely than **gsync**.

2.2 Broadcasting on linear arrays and meshes

The well known spanning binomial tree broadcast algorithm is generalized to obtain several new broadcast algorithms for linear arrays and meshes. These generalizations take advantage of bidirectional communication, the connectivity of two-dimensional meshes, and the difference between node-to-network and network-to-network bandwidth. It is shown how these algorithms can be

further generalized so that any node can be the source of the broadcast message. A partitioning scheme is given that allows these algorithms to be used on linear arrays and meshes of any size. One of these algorithms, the bidirectional spanning tree broadcast, always has lower cost than the recursive halving broadcast for linear arrays. All of these algorithms offer significant performance improvements over the basic spanning tree broadcast. These algorithms do not rely on a knowledge of machine-dependent constants for network bandwidth and latency, so their performance is not as sensitive to changes in machine characteristics as that of hybrid and pipelined algorithms. Performance measurements are given for some of these broadcast algorithms on the Intel Delta mesh.

References

- [1] M. Davis and S. R. Seidel. Global processor synchronization for the Intel iPSC/2 and iPSC/860. Technical Report 92-03, Computer Science Dept., Michigan Technological Univ., May 1992.
- [2] S. R. Seidel. Broadcasting on linear arrays and meshes. Technical Report ORNL-TM-12356, Engineering Physics and Mathematics Division, Oak Ridge National Laboratory, March 1993.
- [3] S. R. Seidel and M. Davis. Global synchronization algorithms for the Intel iPSC/860. *International Journal of High Speed Computation*. (to appear).
- [4] S. R. Seidel and M. Davis. A global synchronization algorithm for the Intel iPSC/860. In *Proceedings of the Scalable High Performance Computing Conference*, pages 220–223, April 1992.
- [5] S. R. Seidel and M. Davis. Global synchronization algorithms for the Intel iPSC/860. Technical Report RNR-92-027, Applied Research Branch, NASA Ames Research Center, August 1992.