

The Next Generation of Hypercube Computers

Trevor Mudge

Dept. Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109

1 Introduction

Massively parallel computers based on hypercube architectures offer an alternative to traditional supercomputers at much less cost. Hypercubes have been discussed in the literature for several decades. As early as the mid-1970's a 256 processor machine was announced by IMS Associates. The processors were Intel 8080's. In 1983 a working hypercube, the 64-node Cosmic Cube, was demonstrated at Caltech [1]. A hypercube of degree n has $N = 2^n$ nodes. The attractiveness of the hypercube over other geometries can be attributed to: 1) the slow (logarithmic) growth in the worst-case internode distance with N ; 2) the slow (also logarithmic) growth in the number of connections to adjacent nodes with N ; 3) the recursive structure of hypercubes—allows multiple users to have disjoint subcubes; 4) the similarity of nodes—there are no special edge nodes as with arrays, for example; and 5) the ease with which trees and meshes of all dimensions can be embedded.

This paper will summarize the developments that one can expect given the expected course of technology in the near future and make the point that software is the major obstacle to the widespread use of hypercube machines.

2 The Present

A number of other machines have been, or are being, developed at Caltech [2,3]. Influenced by this, Intel developed the 128-node iPSC (personal supercomputer) based on 80286/287 nodes. It was the first production machine and was introduced in July 1985. A similar machine, the Ametek System/14 followed.

Currently there are commercial hypercube machines with thousands of processors. The Connection Machine from Thinking Machines has 64K processors arranged in a degree 12 hypercube with each node containing 16 processors [4]. The processors are fairly simple and operate in SIMD mode. The NCUBE/ten has 1024 32 bit processors that operate in MIMD mode. They have a combined potential performance of 500 MFLOPS (million floating point operations per second), although this potential is rarely approached except on special problems. A few thousand processors as complex as state-of-the-art microprocessors is the limit that current technology can support for modest cost air-cooled systems.

The University of Michigan has been a beta site for an NCUBE/ten for the past 6 months and we have had an opportunity to reconstruct the decisions that went into its creation and to perform a preliminary evaluation of the resulting machine [5]. The design process was begun in early '83, and reflects the state-of-the-art at that time. Dominant considerations were packaging constraints, memory integration levels, and custom vlsi integration levels. The result was a node made of only seven components: a custom chip containing the memory interface, internode and I/O channels, and the cpu; and 6 memory chips. The node has a footprint no larger than a playing card. There are 128K bytes of memory per node, and a 10M Hz node can run 1153 Fortran Dhrystones per second (v. 519 for a VAX 11/780) and 445 kilo-Whetstones per second (v. 395 for a VAX 11/780). The chip supports full IEEE 754 floating-point standard arithmetic.

3 The Future

3.1 Hardware

The ability to build reliable inexpensive massively parallel machines has been demonstrated. As a matter of course we can expect faster processors and larger memories. The recent announcement of Intel's iPSC-VX, which includes a high performance vector processing capability at each node, underscores this view. Furthermore, improvements (more speed and higher densities) in designs such as the NCUBE/ten can be expected to keep pace with those in memory chips since each node is dominated by memory.

3.2 Software

Software is a major obstacle. The recent report on the Supercomputer Research Center concludes that the absence of appropriate parallel programming languages and software tools is the single biggest impediment to the successful use of parallel machines [6]. Parallel languages are still in their infancy, and critical development tools such as parallel debuggers are non-existent. This poor support coupled with little understanding of how to express problems as parallel algorithms may limit hypercube machines to special applications. Overcoming this obstacle will be a major challenge.

Acknowledgment This work was supported in part by ARO grant DAAG29-84-K-0070.

4 References

References

- [1] C.L. Seitz, "The cosmic cube," *Comm. ACM*, vol. 28, pp. 22-33, Jan. 1985.

- [2] G. Fox, *The performance of the Caltech hypercube in scientific calculations*, Caltech Report CALT-68-1298, April 1985.
- [3] J.C. Peterson et al., "The Mark III hypercube-ensemble concurrent processor," *Proc. Int'l Conf. on Parallel Processing*, pp. 71-73, Aug. 1985.
- [4] W.D. Hillis, *The Connection Machine*, MIT Press, Cambridge, MA, 1985.
- [5] J.P. Hayes et al., "Architecture of a hypercube supercomputer," *Proc. Int'l Conf. on Parallel Processing*, Aug. 1986.
- [6] Anon, *Report of the Summer Workshop on Parallel Algorithms and Architectures for the Supercomputing Research Center*, Aug. 1985.