

AD-A257 165



1

2

Technical Progress Report 08/01/92 - 11/01/92
Construction of a Connectionist Network Supercomputer
University of California, Berkeley
ONR URI Grant No. N00014-92-J-1617

DTIC
ELECTE
NOV 10 1992
S C D

1 Introduction

This quarter was the second quarter of funding by the ONR for the project. With the project fully staffed we turned our attention to the reexamination of the paper design submitted in the proposal.

Many new ideas arose as we began to work through the details of the system. To limit the time spent on these investigations, we set a goal of early October for a CNS-1 Architecture Specification. The document would be used as an internal design specification. Also, because we were at the point in the project where we have settled major architectural decisions, we wished to get feedback from other researchers in the field. In the second week of October, we distributed the Architecture Specification to approximately 25 experts in the fields of computer architecture and neurocomputing.

With the Architecture Specification written and distributed we held a design review meeting on October 26. The meeting brought together approximately 40 researchers: some part of our team, other colleagues from UC Berkeley and ICSI, and some outside experts. The primary focus of the meeting was the hardware architecture of the CNS-1. At the meeting we discussed many topics; compiling a list of improvements to our design and areas which require further investigation. Also, some of those who were invited and could not attend have agree to provide us feedback at a later date. We will have a round of changes to the current design specification as a result of the work following the design review. We plan to update the Architecture Specification to incorporate these changes and to present a more balanced treatment of the design issues and the architecture subsystems. The revised version will be issued as a technical report around the end of this year.

2 Technical Status

We have made significant technical progress in these three months. Most of these are covered in the enclosed CNS-1 Architecture Specification. We will summarize here the progress in several areas.

The VLSI work reported here is also supported by an NSF experimental systems grant (MIP-8922354) and an NSF PYI award (MIP-8958568).

APPROVED FOR PUBLICATION
Approved for public release
Distribution Unlimited

92-29227



99

2.1 System Packaging

In a high-performance machine such as the CNS-1, physical design and packaging is a major design concern. The physical implementation of the machine has a major impact on its performance, power consumption, and cost. We are at a time when several new chip and system packaging technologies, MCM (multi-chip module), are rapidly being introduced. We have studied how to best exploit these new technologies. In the revised CNS-1 design, we abandoned the traditional computer form of cabinets with backplanes holding large circuit boards for a radically different form. The regularity of a message-passing architecture gives us the ability to choose a different machine form; our requirements for high node density and economic implementation pushed us to exploit this ability. The CNS-1 will be packaged as an upright octagon-shaped tower. The octagon is formed by stacking "hoops" of processors; each hoop consists of 16 identical circuit modules, each containing four Torrent processors with associated memory. The circuit modules are connected to each other using flexible interconnect. This module arrangement naturally implements a cylindrical mesh connection topology. The hoops stack onto an octagonal frame that provides physical support, as well as a wiring and cooling conduit for power and clock distribution and heat exchange.

2.2 Processor Interconnection Network

In addition to the physical implementation of the machine, the processor interconnection network has been the subject of study over these three months. We have deviated from our earlier design of a "ring-of-rings" topology in favor of a more general mesh topology. In our implementation the mesh will wrap around in one dimension forming a *barrel*. There are several reasons for choosing the barrel connection. First, to implement the barrel we will need no other components apart from the processors themselves. This significantly reduces the complexity and physical size of the machine. The barrel has a simple physical mapping with uniformly short connections between neighboring nodes. The barrel is also simple to scale across the size of the machines of interest. The barrel gives very high local bandwidths for those algorithms that map well, including low-level image processing and dense matrix manipulations such as those found in many neural network algorithms. Also, ring topologies map trivially to the barrel, allowing us to carry over existing mapping strategies developed for the RAP processor and our earlier ring-of-rings design. We have studied many of the low level aspects of the communication network using simulation of the CNS-1 network on the Thinking Machine Corporation CM5 machine. This work has resulted in a buffering and routing scheme with much of the advantage of earlier published routing schemes without the associated complexity. However, the details of message level routing are still being studied.

2.3 VLSI design

We have continued a significant effort in the area of VLSI design. We have continued to design and test subsystems for the SPERT processor (the prototype to the Torrent

processor). A important piece of the SPERT processor, the instruction cache, was designed and fabricated. We have just begun its testing. Other detailed design work on the SPERT processor has continued. Also, we performed systematic testing of the SPERT test datapath chip, SQUIRT, and have discovered no new errors. A new post-doctoral visitor, Thomas Schwair, has taken on this responsibility. He will also aid in testability issues on the Torrent processor and the CNS-1 system.

2.4 Software

The October design review concentrated on hardware issues, but also forced the consideration of many software issues as well. These are outlined in the architecture specification document and will be described in detail in the forthcoming technical report, which will have much more detailed software design than previous documents. We also produced a new version (1.8) of the connectionist simulator, ICSIM, with accompanying documentation. This is being tested in our laboratory and at others around the world. Also in this period we completed arrangements for a workshop on software for connectionist supercomputers to be held in Berkeley in April 1993. There are already commitments from the world's leading groups to participate in the workshop.

3 Publications

"CNS-1 Architecture Specification", Connection Network Supercomputer Project, *internal report*, Revision 4.00, October 14, 1992.

John Lazzaro, John Wawrzynek, M. Mahowald, Massimo Sivilotti, and Dave Gillespie, "Silicon Auditory Processors as Computer Peripherals," *IEEE Transactions on Neural Networks*, to appear.

John Wawrzynek, Krste Asanovic, and Nelson Morgan, "The Design of a Neuro-Microprocessor," *IEEE Transactions on Neural Networks*, to appear.

Asanovic, K., Morgan, N., and Wawrzynek, J., "Using Simulations of Reduced Precision Arithmetic to Design a Neuro-Microprocessor" Invited submission to *Journal of VLSI Signal Processing*, to appear in a special issue on Neural Networks, 1992.

Asanovic, A., Beck, J., Kingsbury, B., Kohn, P., Morgan, N., and Wawrzynek, J., "SPERT: A VLIW/SIMD Microprocessor for Artificial Neural Network Computations", *Proceedings of the International Conference on Application Specific Array Processors 1992*.

Asanovic, A., Beck, J., Kingsbury, B., Kohn, P., Morgan, N., and Wawrzynek, J., "SPERT: A VLIW/SIMD Neuro-Microprocessor", *Proceedings of the International Joint Conference on Neural Networks 1992*, volume 2, pages 577-582.

3-4 ✓
fas
Access
Classification

DTIC QUALIFIED 4

By	
Distribution/	
Availability Code	
Dist	Special

A-1