AD-A260 929

DTIC
ELECTE
FEB 1 0 1993
S C D

Technical Progress Report 11/01/92 - 02/01/93
Construction of a Connectionist Network Supercomputer
University of California, Berkeley
ONR URI Grant No. N00014-92-J-1617

1 Introduction

This was the third quarter of funding by the ONR for the CNS-1 project. Our October 26 design review meeting (described in the 11/1/92 quarterly report) and the associated CNS-1 Architecture Specification document subsequently produced two findings with important ramifications for the project.

The first finding is that, in the assessment of our reviewers, the project is technically sound. In particular, there was general agreement that the motivations are reasonable, the physical packaging scheme is plausible, the processor interconnection network is satisfactory, and the VLSI task is attainable. The software strategy was only briefly presented; a more complete software review is planned at a later time. There were several minor points made by the reviewers which are being incorporated in the CNS-1 Architecture Specification.

However, we also received suggestions for making better use of industry and academic "standards", thereby allowing more widespread use of the CNS-1 by others in the field. The subsequent consideration of these suggestion led to the second major result of the review.

We are adopting an industry standard instruction set architecture (ISA) for the scalar processor within the Torrent VLSI chip. Although this decision directly only affects about 15% of the silicon die area (making it slightly more complex to design), the repercussions are extensive in the software area. By using a standard ISA (we have selected the MIPS R3000), we will be able to take advantage of suitable commercially available and public domain system software, tools and application libraries. In addition, machines executing the R3000 instruction set are widely available and will be used as development platforms throughout the life of CNS-1. The basic architecture of the Torrent chip remains the same; it is a scalar processor with a SIMD array of moderate precision datapaths for neural computation. The SIMD array now interfaces to the scalar processor as a vector coprocessor.

93-02436

2 Technical Status

Much of the progress over the past three months is presented in detail in the following documents:

- CNS-1 Architecture Specification (revision 5.0, available in March)
- Torrent Architecture Manual, Revision 1.5 (attached)

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

- Torrent T0 Reference Manual (attached)

A summary of this work is presented below.

2.1 System Packaging

Work continues on refining the system packaging model presented last October—an upright octagonal tower with Torrent modules mounted on the outer surface. Several potential component suppliers have been contacted and a thermal design consultant has been identified. MCM (Multi-Chip Module) technology is being monitored carefully, but the rapid rate of change in the field means the final decision will be delayed until third quarter 1993.

2.2 Processor Interconnection Network

The topology and interconnect schemes for the network have stabilized as discussed in the last quarterly report, namely, a cylindrical mesh with an integrated network interface. Studies are continuing on alternative interface designs, with examination of message passing protocols, buffer sizes and organizations, bandwidth requirements and deadlock avoidance issues. A major extension of our design and simulation studies is required to account for multi-cast message patterns that will characterize much of the traffic on CNS-1. We have begun work on this extension and plan to complete the study in the next reporting period.

2.3 VLSI design

Testing of the SQUIRT datapath chip has continued, with no functional errors identified. The instruction cache test chip has been verified at 20 MHz, the limit of our test equipment. Another test chip with alternative I/O pad designs has been fabricated and tested for performance, power consumption and robustness. This test chip also contains a capacitance measurement structure which has allowed us to refine the simulation parameters for future chips.

A major design effort by our lead processor designer, Krste Asanovic, is reflected in the attached "Torrent Architecture Manual." This document focuses on the vector unit, since the ISA for the the scalar processor is covered in other available documents (e.g., "MIPS RISC Architecture" by Kane and Heinrich). The Torrent architecture specification will serve for a family of processors, with various implementation details described in separate documents.

One such document is "T0 Reference Manual", included in the same binding as the architecture manual. T0 is our name for the first Torrent processor, currently in design. Functionally, T0 will supplant our SPERT design while serving as a "testbed" for the CNS-1 Torrent design. T0 features eight vector processing units, an R3000 compliant scalar

processor, a 128b external data bus which connects to SRAM, and a simplified serial interface based on JTAG. Anticipated peak performance at 50 MHz will be 800 MOPs/second. Tapeout of T0 is planned for Spring of this year.

The second Torrent processor, T1, will be used in the CNS-1. At the recommendation of several of our reviewers, T1 will be designed for 3.3 volt operation. Preliminary analysis of the silicon process parameters suggest that, even in the 0.8 micron technology, a 3.3 volt design will meet our target performance goals while reducing system power significantly.

2.4 Software

The overall software problem for CNS-1 has been greatly simplified by the adoption of the MIPS R3000 ISA. Not only we will be able to port existing compilers, etc., but the development of our own software can be done on MIPS R3000 systems such as DECstations. There are still many difficult problems to solve, but more of the effort can be dedicated to improved performance and flexibility rather than basic support software. We plan to rethink the details of the software plan in the coming quarter, based on several developments. The major software event of the next quarter will be the Workshop on Software for Connectionist Super Computers to be held at ICSI on April 19 and 20. Representatives of all the major groups are planning to participate and we expect a lively discussion. This will also serve as a design review and milestone for our CNS-1 overall software plan.

Meanwhile there has been good progress on software tools and components that will be part of the overall system. The Sather language and its parallel variant (pSather) have advanced rapidly and the designs have stabilized. The ICSIM connectionist simulator was released and Ben Gomes' master's thesis on it was accepted. In the next quarter, Terry Regier will use the simulator in a course and we will begin work on a parallel version for the CM-5, based on pSather.

In a parallel, but gradually converging design effort, the CLONES connectionist simulator has also evolved significantly. CLONES was designed to train very large densely connected networks and has been successfully applied to speech recognition research on the RAP. The new design addresses issues related to parallization of connectionist networks for larger machines such as CNS-1.

Also, significant effort has gone into design and implementation of efficient classes for accessing very large databases of training data. As we expand our experiments with the RAP, we are already running into limitations related to disk size and network loading and reliablity.

Development tools for the T0 chip are currently being designed and implemented. These include an emulator called TREPS (for fast execution on a MIPS based workstation), a simulator (for optimization where cycle counting is important), C++ and Sather support for vector operations, device driver for JTAG, daemon process (for operating system requests by T0), monitor and debugger. The library of matrix and vector routines for T0 is being developed in parallel with the processor design.

3 Publications

Submitted:

Krste Asanović, Beck, J., Feldman, J., Morgan, N., and Wawrzynek, J., "Development of a Connectionist Network Supercomputer," Third International Conference on Microelectronics for Neural Networks, 6-8th April 1993 Edinburgh, Scotland.

To Appear:

Lazzaro, J., & Wawrzynek, J., "Low-Power Silicon Neuron, Axons, and Synapses," chapter in "Silicon Implementation of Pulse Coded Neural Networks," Zaghloul, M.E., Meador, J., Newcomb, R.W., editors, Kluwer Academic Publishers, 1993.

Lazzaro, J., Wawrzynek, J., Mahowald, M., Sivilotti, M., and Gillespie, D. (1993). "Silicon auditory processors as computer peripherals," in *Advances in Neural Information Processing Systems 5*, San Mateo, CA: Morgan Kaufmann Publishers. *to appear Spring 1993 in:* IEEE Transactions on Neural Networks.

Wawrzynek, J., Asonović, K., and Morgan, N., "The Design of a Neuro-Microprocessor," IEEE Transactions on Neural Networks, *to appear Spring 1993.*

Asanović, 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 Spring 1993* in a special issue on Neural Networks, 1993.

DTIC QUALITY INSPECTED 3

Accession For
NTIS [redacted] ☐
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Avail and/or
Dist Special
A-1