High Performance Optoelectronic Neural Network Scheduler

K. J. Symington, A. J. Waddie, T. Yasue, M. R. Taghizadeh and J. F. Snowdon
Dept. Physics, Heriot-Watt University, Riccarton, Edinburgh EH14 5EQ, Scotland.
kjsymington@iee.org, (A.J.Waddie, T.Yasue, M.R.Taghizadeh, J.F.Snowdon)@hw.ac.uk
http://www.optical-computing.co.uk

Abstract: A novel optoelectronic architecture is presented that uses a neural network to provide routing decisions for a packet switch. Recent experiments, combined with simulation studies, show that a digital neuron response improves speed and scalability. The next generation switch (scalable to 62 way) is capable of 2.5 million switch configurations per second. The system includes prioritisation and may be configured for a range of applications.

OCIS codes: (250.0250) Optoelectronics; (200.4700) Optical Neural Systems; (999.9999) Telecommunications Switching.

1. Introduction

This paper discusses an optoelectronic neural network, loosely derived from the Hopfield and Tank model [1], that can be used to rout packets in a packet switch network. Given either a crossbar or banyan switch configuration, the neural network [2, 3] examines the requested output lines of all input packets and selects a setof packets such that mean packet delay is minimised. Since this is a generalisation of the assignment problem other algorithms can be solved such as the TSP, processor load balancing etc.

There are currently two system demonstrators. The first generation[4] was a proof of principle and underlined that such an optoelectronic neural network packet switch construction was feasible. The second generation builds on the first by reducing demonstrator size and improving functionality by the addition of packetprioritisation.

Simulation, employing experimental component characteristics, has been undertaken and has allowed extrapolation of the performance and scaling limitations of this system. Experiment combined with this simulation has also confirmed (perhaps surprisingly given the conclusions in [1]) that digitally thresholded neurons (i.e. digitally driven VCSELs) not only improve the scalability but reduce decision times and hardware requirements.

2. Hardware Description

Figure 1 below shows an overview of the second generation optoelectronic neural network [5].

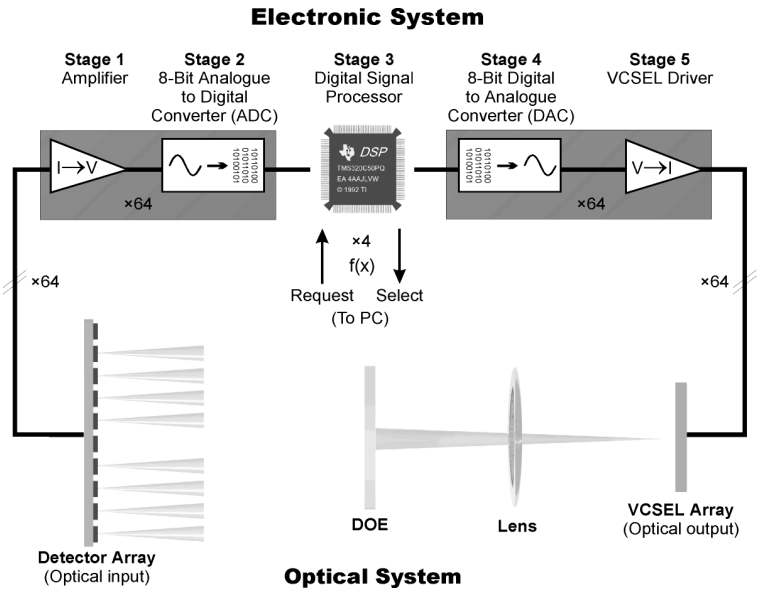

Fig. 1. Optoelectronic neural network system overview. This system is designed to handle an 8 input/8 output switch (N=8).

The neural network consists of a set of neurons, interconnected in an application specific manner, that perform a transfer function on the summation of a set of inputs. Here the optical system is used to perform fixed weight interconnection and summation and the transfer function is calculated by a digital signal processor (DSP).

There are two important measures used throughout this paper. The first is the switch size, (if a crossbar switch is of size N=8 then it has 8 input channels, 8 output channels and N^2=64 neurons) the second is the number of iterations and is a measure of performance. This system iterates until a decision is reached and can only iterate as fast as the slowest component. If the slowest component in the system is capable of 10MHz then 10×10^6 iterations could be performed per second.

The optical system allows scaling of complex neural interconnection strategies in a manner not possible in electronics. There are three components key to the optical system. The VCSEL (vertical cavity surface emitting laser) array is capable of GHz operation and is available in array sizes of 8×8 with bigger arrays currently arriving. The scaling is limited by heat dissipation in larger arrays as the wallplug efficiency of these elements is still fairly low - typically <15%. The DOE (diffractive optic element) forms the interconnection and is capable of very large fan-outs (we have created 128×128 in house). On average 60%-70% of incident power is correctly transmitted by the DOE, the remainder being scattered outside the diffraction window. Uniformity of orders is the DOE system limitation but this is a systematic uncertainty and can be tolerated by the network. The detector arrays are routinely available in large sizes and at GHz speeds.

The electronic system is designed to interface with any packet switch. It requires hardware for each neuron, of which there are N^2 in any system, where each neuron can be considered as five functional stages. Stages one and five simply amplify the signals to and from the inputs and outputs of the optical system. Stage two is an ADC (analogue to digital converter) that converts photocurrent into a digital representation for the DSP that forms stage 3. This part of the system performs the transfer function $f(x)$ and in the current system, one DSP handles 16 channels. Ideally this calculation would be performed by logic on an ASIC or FPGA which would allow each channel to have its own set of dedicated hardware. A DAC (digital to analogue converter) converts the digital signal back to the optical domain which requires conversion of a digitally represented signal to a voltage level which can be done at speeds comparable to the ADC. If digital VCSEL driving is used (see section 3) then there is no need for this component.

Scaling of this system is essentially limited by the size of chip and process technology used. Performance is limited by ADC conversion and transfer calculation times. Speeds of greater than 100MHz are easily attained if custom ASICs are deployed.

3. Digital vs. Analogue

Preliminary design studies indicated that digital operation (with corresponding component savings) might be feasible. Simulation was used to build on our experimental results to determine the relative performance of systems with either analogue or digitally driven VCSELs (corresponding to analogue or digital neurons). Figure 2 plots both types of system.

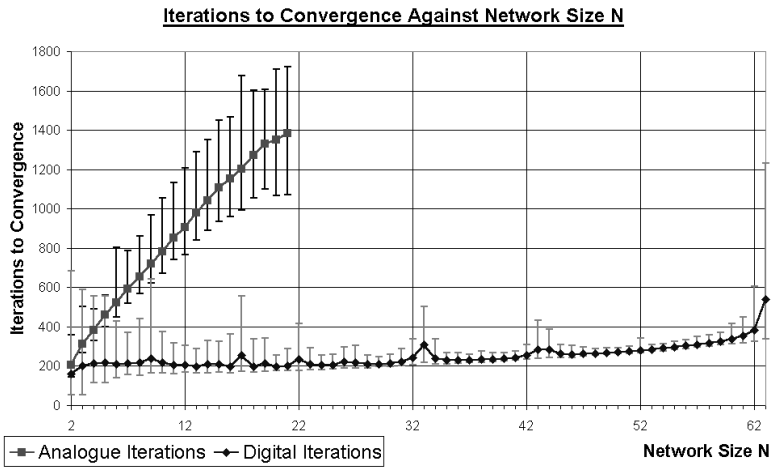

Fig. 2. Convergence in both digital and analogue components.

As network size N increased, the number of iterations required increased linearly in the analogue system. When N became greater than 21 the network started to produce invalid solutions so this was considered to be the limit of analogue scalability. With a digitally driven system, the number of iterations required remained roughly constant over its entire range until the system stared to reach its limit at a size three times greater (N=62) than that of the analogue system. What appears to finally limit our digital solution is ADC discrimination: i.e due to low signal amplitudes it can no longer effectively tell whether a VCSEL is active ornot.

Unfortunately, there is no such thing as a free lunch. Further simulation revealed that solution optimality is slightly degraded in the digital system at low load levels (figure 3). This is due to an impetus being added by digitally thresholding that can cause neurons to converge a little too hastily. Thus a sub-optimal solution of N-1 neurons *on* (equivalent to packets routed) rather than the optimal N will be produced 9% of the time rather than 3% of the time with an analogue system.

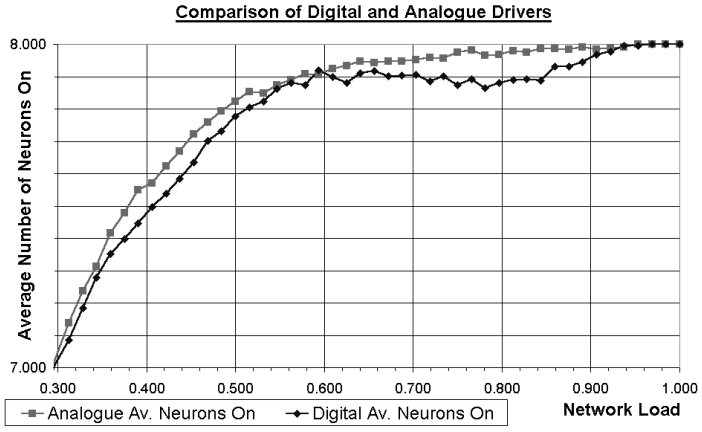

Fig. 3. Comparison of quality of solution between digital and analogue systems.

However, a digital implementation removes a potential bottleneck from the system by making digital to analogue conversion unnecessary. Removing hardware to increase performance and scalability is very tempting even with the disadvantage of a slightly less optimal solution.

4. Conclusions

Performance wise, a 100MHz system is easily feasible, with up to 1GHz a possibility: the bottleneck is calculation of the transfer function and ADC conversion times. Scalability is limited by VCSEL array size which is currently N=16. This problem could be circumvented by using multiple quantum well modulators (MQW) in which case current DOE non-uniformities becomes the next limiting factor. Using a digital system we can therefore foresee a system that can provide 2.5 million switch configurations per second regardless of whether the switch has 2 or 62 channels.

An interesting piece of further work involves the support of variable length packets. An examination of the quality of service (QoS) that is provided by both this and the current system is required to underline the system's performance at higher load levels.

5. References

[1] J. J. Hopfield and D. W. Tank, "'Neural' Computation of Decisions in Optimisation Problems", Biological Cybernetics, **52**, pp. 141-152 (1985).

[2] P. W. Protzel, D. L. Palumbo and M. K. Arras, "Performance and Fault-Tolerance of Neural Networks for Optimisation", IEEE Transactions on Neural Networks, **4**, 4, pp. 600-614 (July 1993).

[3] J. Ghosh, A. Hukkoo and A. Varma, "Neural Networks for Fast Arbitration and Switching Noise Reduction in Large Crossbars", IEEE Transactions on Circuits and Systems, **38**, 8, pp. 895-904 (August 1991).

[4] R. P. Webb, A. J. Waddie, K. J. Symington, M. R. Taghizadeh and J. F. Snowdon, "Optoelectronic Neural-Network Scheduler for Packet Switches", Applied Optics, **39**, 5, pp. 788-795 (February 2000).

[5] OSA: A. J. Waddie, K. J. Symington, M. R. Taghizadeh and J. F. Snowdon, "An Optoelectronic Neural Network: Implementation and Operation" Optics in Computing, Quebec City, Canada, R. A. Lessard and T. Galsitan, Vol. 4089 of OSA Proceedings Series (Optical Society of America, Washington, D.C., 2000), pp. 304-310.