High Bandwidth Dynamically Reconfigurable Architectures using Optical Interconnects

Keith J. Symington¹, John F. Snowdon¹ and Heiko Schroeder²

¹ Department of Physics, Heriot-Watt University, Edinburgh U.K. kjsymington@iee.org and J.F.Snowdon@hw.ac.uk ² Department of Computer Science, Loughborough University, Loughborough, UK.

Abstract. Optoelectronic interconnects are one means of alleviating the ever growing communications bottlenecks associated with silicon electronics. In chip-to-chip and board-to-board interconnection, the bandwidths presently (if experimentally) available far outstrip what is predicted possible in electronics until into the next decade. Such high bandwidth possibilities demand a rethink of conventional computer architectures where bandwidth is always at a premium. The combination of dynamic reconfiguration in electronics with this new technology may enable a new generation of architectures.

1 Introduction

Dynamically reconfigurable field programmable gate arrays (FPGAs) combine, in principle, the speed of dedicated hardware with the flexibility of software. As with all VLSI systems, the increasing density and speed of silicon circuits frequently transfers the performance bottleneck of a system to its communications. Optoelectronic interconnects are widely considered as a potential solution to this problem and are already being deployed (at a crude level) in commercial systems. Aside from the potential of optics for raw data throughput (unattainable in conventional systems), what is perhaps more exciting is the enabling of new architectural concepts by a combination of this throughput with the potential to reconfigure at high speed. This paper begins with a discussion of the potential for free space optoelectronic interconnects and a description of the types of hardware currently available. Three optoelectronic systems are then described; a dedicated digital sorter, a neural network switch controller and a generic parallel processing interconnect harness with a role for FPGAs suggested in each case.

2 Optoelectronic Interconnects

As the length or density of electrical interconnection increases, it suffers from increased wire resistance, residual wire capacitance, fringing fields and cross-talk. The maximum bandwidth of electrical systems has been independently estimated [1] as being around $500A/L^2$ THz (where A/L^2 is known as the aspect ratio), which for a 10cm off chip connection is around 150GHz. Free space optics can deliver far higher bandwidths. The free space optical systems considered here are constructed by flipchipping optical detectors and emitters (or modulators) onto silicon circuitry (see section 3) so the whole ensemble can be viewed as "conventional" silicon equipped with a large number of "optical pins" normal to the chip surface. The advantages of such a system are as follows

Off Chip Data Rates: The number of optical pins that can be driven depends primarily on thermal and real estate considerations. Currently we can drive 4,096 channels from 1cm^2 and see no real obstacle to reaching >10,000 channels. These channels can be driven at speeds much greater than that of the CMOS and take considerably less power to drive than pads and wire bonds. Of course one may still connect conventionally to the chip as well. The sorting demonstrator system described in section 4 has an off chip data rate of 200 Gbs⁻¹ which is what will be required by the semiconductor industry (according to the SIA roadmap) in 2007.

Bandwidth in "Busses": The bandwidth sustainable in a free space optical relay (see section 4) is far higher than that in an electrical bus. A 1 cm^2 relay can carry >100,000 channels. Currently we are driving at 200MHz (CMOS limited) giving a bandwidth of approximately 20 Tbs⁻¹. Devices may routinely be driven at 10 Gbs⁻¹ so the relay can handle 1,000 Tbs⁻¹ if we are not CMOS limited (the theoretical limit is actually much higher).

Distance: Optical signals (both guided and free space) are attenuated far less than electrical. Optical lengths of the order of meters are attainable without a significant increase in driving power.

Non-Local Interconnections: The non-interacting nature of free space optical channels means that they can pass through each other to form any desired interconnection topology without cross-talk. Interconnects such as the perfect shuffle and the hypercube thus become relatively simple to implement. This also reduces skew as very large variations in wire length can be avoided.

Data Acquisition: The naturally parallel nature of the connections implies high parallelism around the machine: e.g. to and from memory and peripherals.

Reconfiguration: It is also possible to reconfigure optical relays in the optical domain. This is not considered here.

3 Integrating FPGAs with Optoelectronics

The novel components may be thought of as consisting of three basic stages. The first stage, the input stage, consists of a detector array that is capable of receiving digital optical input. The second stage is the processing stage and consists of a dynamically reconfigurable FPGA system that could be considered as one or more configurable logic blocks (CLBs) corresponding to a single detector input. The final stage is the output stage and consists of an optical emitter or modulator, again corre-

sponding to one or more CLBs. This combination will be referred to as an Optical FPGA (OFPGA).

The OFPGA (figure 1) is capable of communicating internally electronically, or with any other local electronics for that matter, and can be viewed as a standard dynamically reconfigurable FPGA with an extra optical input and output available to a specific CLB or set of CLBs. Any optical input has the potential to reprogram a CLB by interlacing configuration information using a predefined protocol, or to reprogram another CLB in another system by interlacing the same configuration information onto its optical output stream.



Figure 1: Combination of detector, FPGA and VCSEL arrays.

In general the optoelectronic interface is

constructed of hybrid processing chip technologies, which employ GaAs optical chips hosting detectors and emitters, flip-chipped [2] on top of the Si FPGA. The combination of input, processing and output elements is generally known as a smart pixel. Figure 1 shows a three layer chip which makes visualisation of circuits simpler (actual components at present are two layer with input and output on the same surface).

4 Optoelectronic Dynamically Reconfigurable Systems

The Optoelectronic Sorting Demonstrator [3]: The architecture of the demonstrator utilises optoelectronics as described above and exploits a non-local interconnect, in

this case the perfect shuffle. Figure 2 shows a schematic of the sorting demonstrator. Each OFPGA array can convert electrical or optical inputs into electrical or optical outputs. The data to be sorted are entered sequentially into the processing loop through electrical I/O as shown. Sixteen bit planes of 32x32 bits (the number of optical communication channels) may be entered in this version. At run time, the 2D perfect



Figure 2: Optoelectronic sorting demonstrator modified to include inline components.

shuffle is performed by a lens operation during each cycle of the machine and all computations are performed in parallel by the FPGA. The total number of cycles scales as $(\log N)^2$ for a Batcher's bitonic sort of *N* data points. At the conclusion of the computation, the sorted set of data can be sequentially downloaded to the electronic domain.

We have described a sorting system built using two OFPGAs; however it is the uses of the FPGAs reconfigurable aspects that are of interest here. A first level of flexibility could be introduced in that the CLBs could be reconfigured to optimise execution time for different data array sizes [4]. As a further level, the iterative perfect shuffle shown here forms (with suitable node switching) an omega class network capable of arbitrarily permuting data. Bearing this in mind, the set up could be used to implement any algorithm or multistage switching function given the right logic configuration. Indeed it is known that many algorithms map exceedingly efficiently onto this topology (the FFT being the classic example) and performance results at this level of parallelism would be of some interest to the engineering community. Therefore the combination of the FPGA and the high bandwidth optical interconnect gives us a general purpose fine grained massively parallel processor.

The Neural Network Packet Switch Controller [5]: Although the physical layout of the neural switch controller is somewhat different to the sorting module described above, from an architectural perspective the neural network may be formed from the module merely by replacing the optics that are used to generate the perfect shuffle with a diffractive optical element (DOE) [6] such as that shown in figure 3. The effect of the element is to fan-out in a space invariant manner every input channel into the pattern shown. This may readily be seen to effectively amount to an analogue sum over rows and columns of the outputs of the previous in-line device. The system as it stands at present has a fixed set of weights and is designed to perform this

single task. The inclusion of the FPGA stages will allow, in the first instance, programmed weights to be considered which will allow the network to be configured for a wider variety of tasks such as the travelling salesman problem or other optimisation problems. As in the above example however, the most exciting possibility is being able to reconfigure these weights in near or at real time so that fully adaptive, supervised and unsupervised learning schemes may be implemented.



Figure 3: Illustration of a single input beam imaged onto multiple detectors using a DOE.

So the combination gives us a neural network with fully real time adaptive weights with an interconnection density that could never be achieved electronically.

The Generic Multiprocessor Harness: The concept of optical highways [7] is of perhaps the most interest in that a general purpose multiprocessor solution can be envisaged. Figure 4 schematically shows a highway used to connect nodes in an

arbitrary topology, with each node having access to >1,000 channels. The interconnect is point to point and hard wired, with several thousand channels being passed from each node via a smart pixel interface into a free space optical relay system which can hold (under reasonable assumptions regarding aberrations) several hundred thousand channels. Polarising optics are used to determine the position of every individual interconnection link thus defining the computational topology.



Figure 4: Optical highway for connection of multiple processing nodes.

The data used in this design is extrapolated directly from the components used in the sorting demonstrator (number of optical channels off chip is 2,500, clock speed 250MHz, data width is 64 bits) which shows the highway to support bisection bandwidths in excess of 10,000 Tbs⁻¹ connecting up to 1 million nodes (depending on topology) [7]. Each node might consist of a microprocessor, memory and cache, and one or more OFPGAs to handle communications.

The role of OFPGAs in this architecture is to optimise the communications and processing in real time during execution of a range of algorithms. The bandwidth of the communications in this system is sufficiently high to implement a flat memory model, but superior performance on particular algorithms may be obtained by changing topology in the interconnect harness. In essence, the destination of any data channel output into the optical domain is determined by the spatial location of the emitter output upon the OFPGA. Thus by re-routing signals within the OFPGAs, a particular global topology may be established. The machine could of course be configured arbitrarily into several differently connected regions if this was desirable. In addition, the functionality of the interface may be changed, for example, OFGAs allow us to configure the interface as a router (necessary if we wished to utilise a hypercube in the optical domain) or as a simple switch of high throughput (necessary if we wished to use a large crossbar in the optical domain). For nodes of sufficient complexity, the maximum throughput of the system may often be attained by changing the width of data words as well as the topology so as to keep all communication channels busy. OFPGAs could support variable width multiplexing as well as routing at the interfaces.

For the generic interconnect the combination of FPGAs with high bandwidth optoelectronics enables an intelligent communications interface to be constructed which allows the maximisation of the ultra high throughput available. In turn this enables real time optimisation and load balancing of the whole machine over a range of computational models.

5. Discussion

Free space optical interconnects appear to offer a tremendous enhancement in all fields of computation and raise interesting computer science and architecture issues. There are many engineering issues to be confronted before such interconnects can be routinely deployed. Alignment tolerances both at set up and during operation remain a difficult problem outside of the laboratory environment. The main approaches followed today include rigid passive integrated structures or active alignment techniques [8]. Although such techniques may appear difficult and expensive, one must remember that something as cheap and commonplace as a CD player uses precisely these adaptive optics to maintain spot tracking even on portable machines!

This paper has described free space interconnect techniques for the non-specialist reader and qualitatively considered the possibility of combining FPGAs with optoelectronics. Three novel systems have been considered in which the combination of reconfigurability with the high bandwidths and interconnectivities offered by optics enable new architectures.

References

- Miller, D. A. B., Ozaktas, H. M.: Limit to the Bit-Rate Capacity of Electrical Interconnects from the Aspect Ratio of the System Architecture, Journal Of Parallel And Distributed Computing, 41, No. 1, pp. 42–52, (1997).
- Makiuchi, M., Hamaguchi, H., Kumai, T., Aoki, O., Oikawa, Y., Wada, O.: GaInAs pin Photodiode/GaAs Preamplifier Photoreceiver for Gigabit-rate Communications Systems using Flip-Chip Bonding Techniques, Electronics Letters, volume 24, number 16, pages 995-996, (August 4th 1988).
- Gourlay, J., Yang, T., Dines, J. A. B, Snowdon, J. F., Walker, A. C.: Development of Free-Space Digital Optics in Computing, Computer, 31, 2, pp. 38-44, (1998).
- 4. Akl, S.G.: Parallel sorting algorithms, Academic Press, (1985).
- Webb, R.P., Waddie, A.J., Symington, K.J., Taghizadeh, M.R., Snowdon, J.F.: An Optoelectronic Neural Network Scheduler for Packet Switches, Submitted to Applied Optics, (May 1999).
- Taghizadeh, M., Turunen, J.: Synthetic Diffractive Elements for Optical Interconnection, Optical Computing and Processing, 2 (4), pp. 221-242, (1992).
- Dines, J.A.B; Snowdon, J.F; Desmulliez, M.P.Y; Barsky, D.B; Shafarenko, A.V., Jesshope, C.R.: Optical interconnectivity in a scalable data-parallel system, Journal of Parallel and Distributed Computing, 41, pp. 120-130, (1997).
- Yang, T. Y., Gourlay, J., Walker, A. C.: Adaptive Alignment with 6-Degrees of Freedom in Free-Space Optoelectronic Interconnects, OSA Technical Digest, Optics in Computing, pp. 8-10, Snowmass, (April 1999).