Low latency optoelectronic processor-memory interconnection demonstrator

Stuart Fancey[*a], Jurgen Jahns[*a], Paul Lukowicz[**b], Janusz Grzyb[**b]
[a]Optical Information Technology, University of Hagen, Germany
[b]Electronics Laboratory, ETH Zürich, Switzerland

ABSTRACT

A principal performance limitation of current computers is memory access latency. The random access time of DRAM can be as low as 20ns but the overhead imposed by communication latency can increase the retrieval time to 150ns in single processor systems or 1ms in large multiprocessor systems. Optically interconnected VLSI offers the possibility of reductions in the communication component of memory latency of an order of magnitude. The improvement arises from the potential of direct high bandwidth low-latency links between any one chip and each one of a set of others. This potential principally arises from the ease of an optical implementation of fan-out and fan-in operations, together with the intrinsically high bandwidth of optical links. We have designed a scaleable system of processor-memory interconnections to explore this technology. Optical fan-out and fan-in modules will link a single processor to a bank of memory chips. The approach allows for multiple processors to be connected to multiple memory banks in an analogous fashion. The demonstrator will use 1-D VCSEL and photodiode arrays to provide optical i/o for the CPU and memory chips. The optical fan-out, fan-in and image relay can be implemented using an integrated planar optical system.

Keywords: Optoelectronics, Optical Interconnects, Parallel processing, Free-space Optics, Multi-processor systems

1. INTRODUCTION

Recent interest in the application of free-space optics to VLSI i/o has been largely driven by the approaching aspect ratio limits of conventional (2-D) electrical interconnects[1]. The limitations of electronics might be most readily observed in systems whose interconnections combine a greater or lesser degree of parallelism, a need for single channel data rates in the Gbyte/s regime and a potential for communication distances in the range of centimeters to meters. The memory access system of a multi-processor computer has just this set of attributes.

In this paper we will discuss memory access latency and explore the degree to which free-space interconnections might be expected to make a dramatic impact on this parameter. We will then introduce our work on the implementation of a parallel memory access scheme based on planar integrated free-space optics. The design and fabrication of such a planar optical element will be discussed, along with the crucial question of integration of free-space optics with the world of chips, fiber-ribbons and printed circuit boards.

2. THE MEMORY ACCESS LATENCY PROBLEM

Memory latency is among the most critical performance factors in today's computers[2,3]. Modern microprocessors can execute several instructions in 1 nanosecond. At the same time main memory access typically takes between 100 and 200ns. In larger multiprocessors it can even take up to several hundred nanoseconds. High memory latency is often attributed to the high cost of fast (but low-density) SRAM memory compared to slow (but high-density) DRAM components. While true, this is relatively unimportant since DRAM memory modules with a latency of 22.5ns and SRAM modules with a latency of less than 10ns are available. A close look at today's memory systems shows that the really important factor in main memory latency is the communication delay. A memory system in a multiprocessor has to interconnect several processors, one or more i/o controllers, and tens of memory chips distributed over a large printed circuit board (PCB). In large parallel machines the interconnection might even have to span several PCBs or cabinets. Each

[*] Corresponding author: Stuart.Fancey@FernUni-Hagen.de; phone +49 2331 987 1123; fax +49 2331 987 352; http://www.fernuni-hagen.de/ONT/LGONThome.html; Optical Information Technology (ONT), University of Hagen, Feithstr. 140 / PRG, 58084 Hagen, Germany. [**] lukowicz@ife.ee.ethz.ch; phone :+41 1 6327614; fax +41 1 6321210; http://www.ife.ee.ethz.ch/~lukowicz; Electronics Laboratory, ETH-Zürich, ETZ, H93, Gloriastr 35, 8092 Zürich, Switzerland.

processor needs a bandwidth in the range of 1GByte/s and the ability to broadcast certain operations to all other processors and i/o controllers. The problems with such an interconnection stem from the sensitivity of high bandwidth, low latency electrical interconnections to length, fanout and layout geometry. While very fast links are possible between neighboring chips (e.g. between a processor and an off-chip cache), they are unfeasible in large, complex, systems like the processor memory interconnection. Instead most memory systems are implemented in several stages, each using many fairly low frequency (100-200MHz), high latency links to achieve the required bandwidth.

In addition to the (cache) memory hierarchy just described, the memory access latency problem has been approached, to date, by the adoption of prefetching algorithms, or the implementation of out of order execution and speculation. These approaches give mixed results as they depend strongly on the pattern of memory access. In addition, they scale poorly with processor speed increases since they involve intensive table manipulation and housekeeping.

Let us look at an example of the complexity of a symmetric multiprocessor. A typical multiprocessor (Sun UltraSPARC) layout is illustrated in Figure 1. The two CPU's in such a system are connected to 16 memory SIMMS. This connection requires two i/o controllers (U2S and FFB), a system controller (DSC), an electronic crossbar (XB1) and further interface stages (UDB and CBT). It can be seen that, since each of the communication steps (crossbars, contention resolution, protocol handling etc) introduces 10-20ns of latency, the total memory access latency of the memory access system can easily exceed 100ns. Our work is based on the argument that the intermediate stages contained within the shaded area of Figure 1 are replaceable by a new, and simpler, access scheme based on free-space optics.

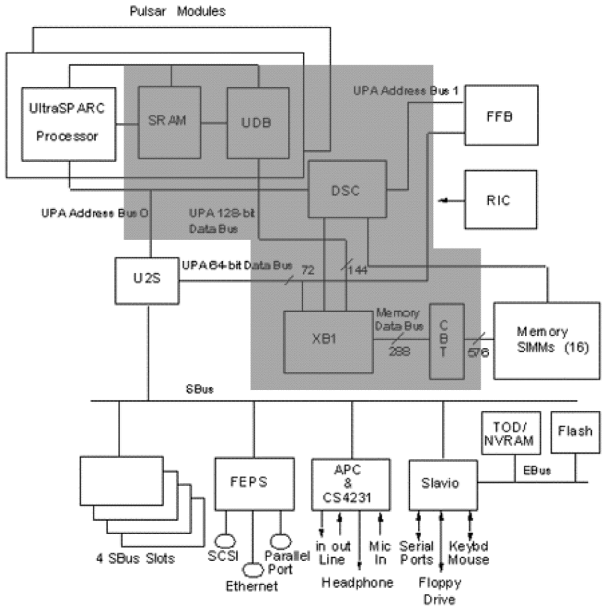

Figure 1: Schematic of an Ultra SPARC workstation showing intermediate steps replaceable by optics

3. PLANAR OPTICS

Before looking in more detail at the access scheme we propose, we will introduce the packaging technology that we believe is a candidate for the integration of free-space optics with the electronic domain, in at least some contexts. Planar integrated free-space optics[4,5] (planar optics) is, in essence, the folding of optical paths within a transparent substrate. The optical elements which make up the system are positioned on the surfaces of the substrate. By adopting a planar configuration, industry-standard fabrication and assembly processes are immediately made available. Photolithography, evaporation, etching, flip-chip bonding, ball bonding etc are all applicable. More explicitly, planar optical systems use refractive or diffractive lens arrays, evaporated metal and thin-film mirrors, diffractive gratings, refractive and diffractive prisms, VCSEL arrays attached by flip-chip bonding and so on. It is immediately clear that, by using lithography and similar fabrication processes, the alignment precision required for free-space interconnection is immediately satisfied. In addition, potential replication is considered by borrowing semiconductor industry processes. Further benefits include the alignment stability that a common substrate provides, the compactness of a folded mm-scale module and the ease of integration of a flat, insulating substrate.

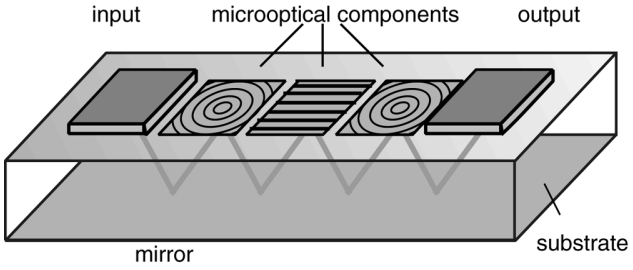

Figure 2: The planar optical principle

The principle is illustrated in Figure 2. The glass or plastic optical substrate (signal layer) is some 3-10mm thick and contains the propagating signal beams as shown. Passive optical components are defined on the outer surfaces to direct the beams and active optoelectronic chips provide opto-electric conversion. A more detailed illustration is show in Figure 3

where we can also see the integration of fiber (-ribbon) coupling to and from remote devices and the direct integration of electrical lines on the substrate allowing it to act as a mechanical mount for purely electronic devices and allowing (electrical) connection between (opto)electronic chips on the optical substrate and chips on a carrier PCB or remote board. An example of a planar optical substrate is shown in Figure 4. In this photograph we can see three circular diffractive lens elements, two square mirrors and two lighter-shaded square grating couplers that make up this system. The lens and mirror regions have been coated with aluminum in this case to provide the optical confinement.

We will look more closely at the fabrication and integration technologies in later sections but, having indicated the form of optical interconnections that we propose, we can return to their possible application in a memory access scheme for parallel computing.

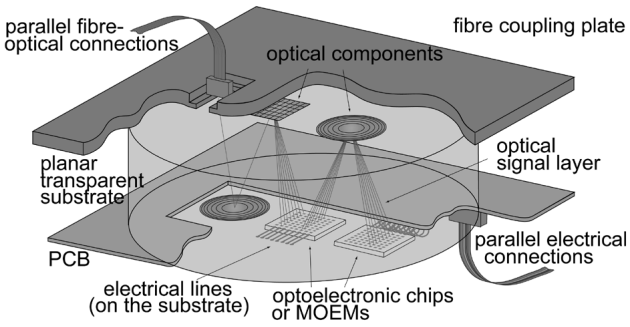

Figure 3: An example of a fictitious planar optical system, illustrating the packaging and integration technologies under investigation

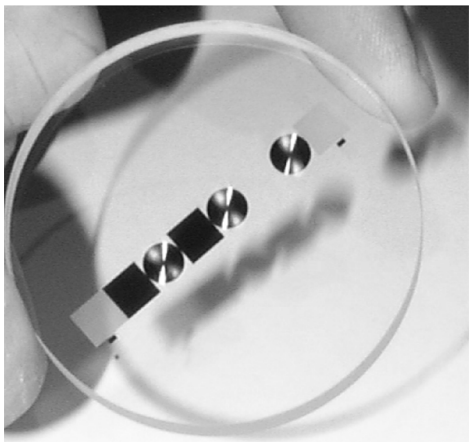

Figure 4: A photograph of a planar optical system on a 2 inch diameter quartz glass substrate

4. A POTENTIAL MEMORY ACCESS ARCHITECTURE

The demonstration system we are building is based on the GigaCache architecture[6]. The architecture is shown, in a simplified form for clarity, in Figure 5. In the figure, two processors are shown connected to one of N_G memory bank groups by direct optical links. The individual memory bank groups contain N_B memory banks, which consist of N_C RAM chips electrically connected to a controller chip. Individual optical channels are established between each processor and each memory bank group. Within the memory bank groups, the messages from the processors are broadcast to the individual memory banks using an optical fan-out. Data flowing back to the processors from a memory bank group is controlled by the memory bank controllers within that group to prevent conflict in accessing the channels to the processors: only one memory bank may return data from within a given memory bank group, at any one time.

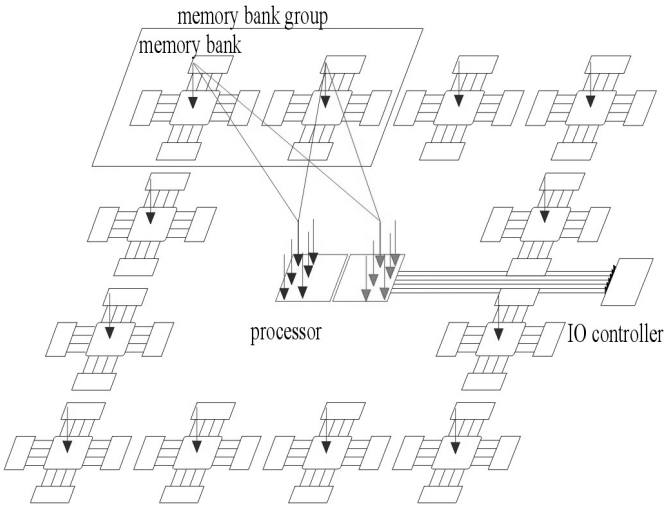

Figure 5: An illustration of the GigaCache architecture with connections from two processors to one of N_G memory bank groups.

We can look more closely at the optical functional unit of this proposed architecture in the diagram Figure 6. In this figure, a memory bank group containing two memory banks of four RAM chips, is shown to have a bi-directional optical link to each of two processors. Within the dotted line on the Figure, we see the components that are bound together within, and by, a planar integrated optics based multichip module. We shall now look more closely at the function of this optical sub-unit before exploring technological aspects of the fabrication and integration.

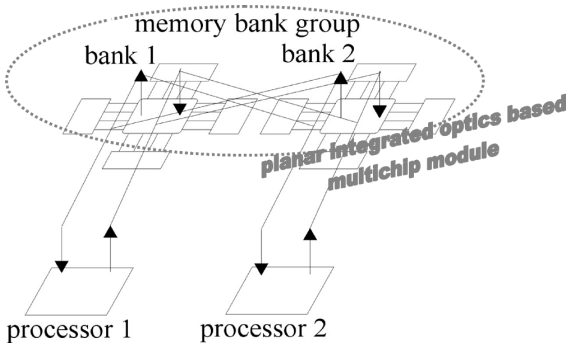

Figure 6: The functional partition of the GigaCache architecture showing the optics-based sub-unit.

The important point to notice about such a building block of free-space optical functionality is its simplicity. The function is a 1:N_B fan-out from a fiber ribbon to the memory bank controller i/o and a N_B:1 fan-in from these memory bank controllers to the fiber ribbon (see Figure 7). That fiber ribbon might itself lead to the central port of an identical fan-out/in module where the place of the memory i/o's is taken by the i/o's from multiple processors. The extension of this scheme to multiple processors and memory banks is illustrated in Figure 8 where, using a basic fan-out/in element as shown in Figure 7 we can see that four processors can be connected to four memory groups, each containing four memory banks, all with only one layer of electronics (the memory controllers local to the memory banks) to introduce memory access latency.

The aim of our project is to implement such an architecture in, initially, a one or two processor system. The memory access latency improvement we expect is significant. With a two processor system connecting to 256-1024Mb of ESDRAM we would hope to see memory access latency in the region of 20-40ns ie approaching an order of magnitude improvement. Our system will be based on ARM processor architectures, running an adapted Linux symmetric multiprocessor kernel.

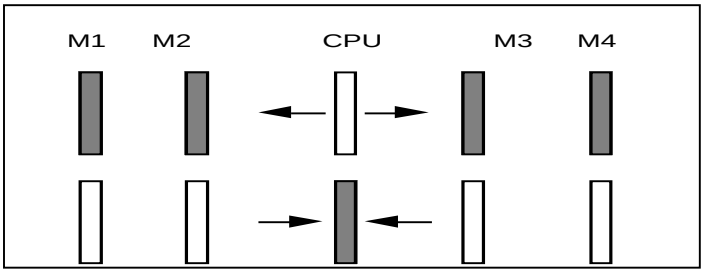

Figure 7: A schematic of the optical function within a memory bank group. The external fibre i/o to one or more CPUs is labelled "CPU" and the fan out(in) to(from) four memory banks "M1"..."M4" is shown

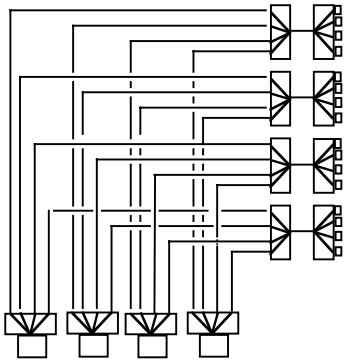

Figure 8: Schematic of potential multi-processor scheme built from repeated fan-out/in modules of the kind illustrated in Figure 7

5. TECHNOLOGY AND PACKAGING

It has been stated above that planar optics is free-space optics in a folded package. It must be stressed that it is not in any way a waveguide configuration but is truly free-space. The design processes therefore follow closely those that are used for more conventional optical systems. We will examine the fan-in direction of our optical module in more detail though much of what we will present is applicable to the fan-out direction also.

The interconnection problem that we need to solve here is a typical dilute image situation, that is to say that we are imaging widely separated discrete areas (10μm diameter VCSEL apertures) onto other discrete areas (multi-mode fiber cores). There is extensive justification in the literature for a hybrid imaging solution to such a dilute imaging problem[7]. Hybrid imaging is a combination of micro-lenses and bulk lenses used to segment the space-bandwidth product in a way that gives low aberrations. An array of fast microlenses collimates the VCSEL output and relatively slow bulk lenses relay the image to a second array of microlenses that focus the beams into the target areas. There are various configurations of such hybrid imaging but the one that we have adopted is the so-called lens-waveguide approach where intermediate bulk lenses increase the resiliance of the system to external perturbations[8].

Let us look at the system in more detail, in Figure 9 below. The side-view in the top diagram shows the fan-in from two inputs to the fibre output. The other two inputs are symmetrically arranged and are not shown. The two top-views below show the two separate imaging systems for the outer and inner VCSEL arrays respectively. Note that only four channels are shown out of the real twelve, for clarity. Microlens arrays *1* (and *6*) perform the collimation of the VCSEL emission, together with introducing the propagation angle required. The bulk lenses *2,3,4* (*7,8,9*) act as the lens relay system, relaying the image to a second set of microlenses 5 (*10*) which perform the imaging and, in combination, the fan-in. Note that the overall length of the optical system is around 10mm

In our demonstration system we will implement the lenses and microlenses using diffractive optical elements. The design freedom of diffractive elements can be used to further optimize the lenses for off axis use[9]. We will use a simple aluminum mirror coating to provide the reflectivity for confinement of the optical beams. This leads us on to questions of power budget. In the fan-out path of our system we must be aware of the power budget implications. In the fan-out optical path we would predict a total transfer efficiency of only 1% and this is the limiting factor for larger degrees of fan-out, to implement

such systems as we saw in Figure 8 above. A key improvement will be the adoption of better mirrors. Aluminum, though convenient, has a relatively unsatisfactory reflectivity (around 85%) at the commonly used VCSEL wavelength of 850nm. We are also working on the fabrication of more general refractive micro-optical profiles to further improve efficiency.

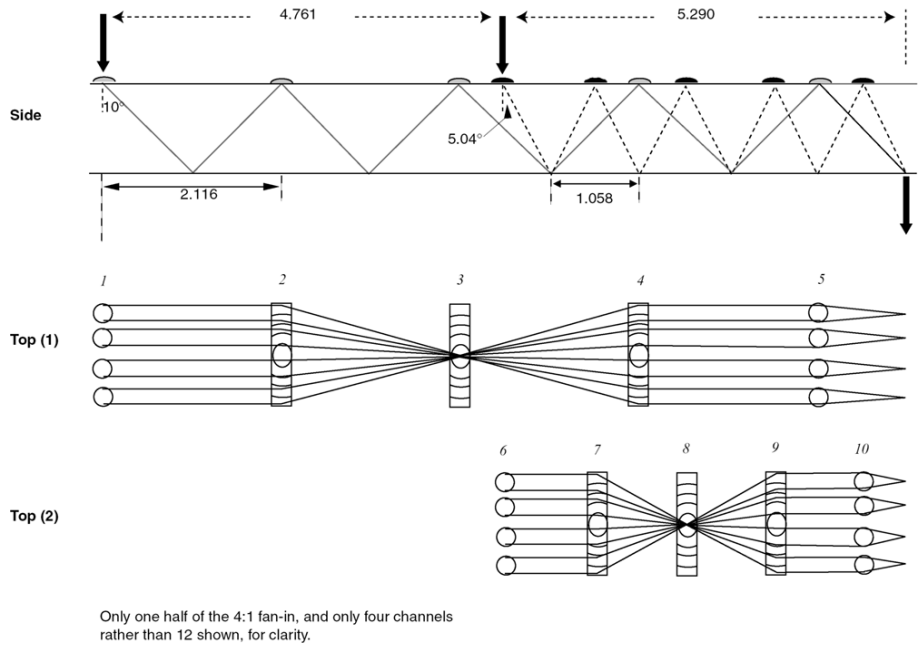

Figure 9: Optical scheme of fan-in module. Dimensions in mm unless stated.

The lithographic processes used to define the planar optical components provide sub-micron precision but the question of tolerances is still important since stability of alignment during temperature (or VCSEL wavelength) fluctuations is desirable. Studies have been performed of the lens-waveguide design of planar optical interconnect[8,10] and it has been shown that tolerance to deviations in wavelength, substrate thickness (related to temperature) and wedge angle (related to poor manufacture) are all considerably improved by adoption of this tolerant configuration over a conventional 4-F hybrid imaging scheme.

A key element in our planned demonstration of free-space processor-memory interconnection is the degree of integration that we can include. The direct attachment of active optoelectronic devices (VCSEL arrays, detector arrays) to the optical substrate is an important example of integration as the alignment of these devices to the optical system is critical, as is their integration with external electrical supplies of data, power, control signals etc[11]. It is, furthermore, of interest to explore the use of a plastic or glass optical substrate as a mechanical mount for electronics, to replace the PCB or to act as a communications layer behind the PCB. In our present work we do not intend to go so far as to replace the PCB entirely but we do aim to construct a module, a form of multi-chip module, that can be integrated mechanically and electrically with the PCB using standard ball-bonding techniques. The module itself will contain the planar optical substrate, on which will be flip-chip bonded VCSEL and detector arrays, together with driver and receiver chips. The electrical interconnection between these, and their electrical connection to the ball-bonding pads leading onto the host PCB, will be implemented using evaporated co-planar lines on the glass. We have performed

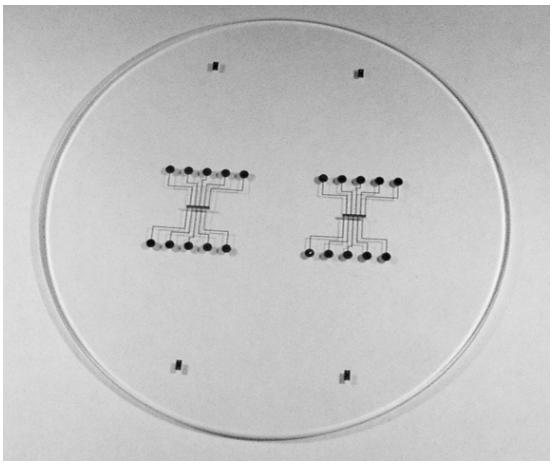

Figure 10: An example of sequential flip-chip bonding of VCSELs to glass with electrical contact lines leading to bond pads.

experiments with sequential flip-chip bonding of VCSEL chips onto glass in close proximity and their connection to bond pads using evaporated silver or gold lines. A photograph of such an experimental result is shown in Figure 10.

The attachment of VCSELs and other heat-generating chips (receivers, for example) onto a flat glass or, worse, plastic substrate raises questions of thermal management. The use of wavelength-critical diffractive optics in the system requires a degree of predictability and stability in the equilibrium VCSEL operating temperature, for example. In addition to the adoption of tolerant imaging schemes (see above), work has been performed to develop packaging techniques to integrate thermal vias with the electrical lines[12], and to integrate the VCSEL devices into heat-spreading solder structures[13]. This approach requires further work to prepare systems for the wide-range of ambient temperatures encountered outside the controlled conditions of the laboratory.

Another important integration step, crucial for the concept of a memory access scheme for (distributed) parallel computing, is the integration of free-space optics in the planar regime with arrays of optical fibers. As part of a related project[14], a technique was devised to fabricate an interface plate for standard 12-way fiber MT-connectors. The interface plate can be aligned with the optical substrate, allowing free-space to fiber coupling. This interconnection optomechanics is shown in Figure 11.

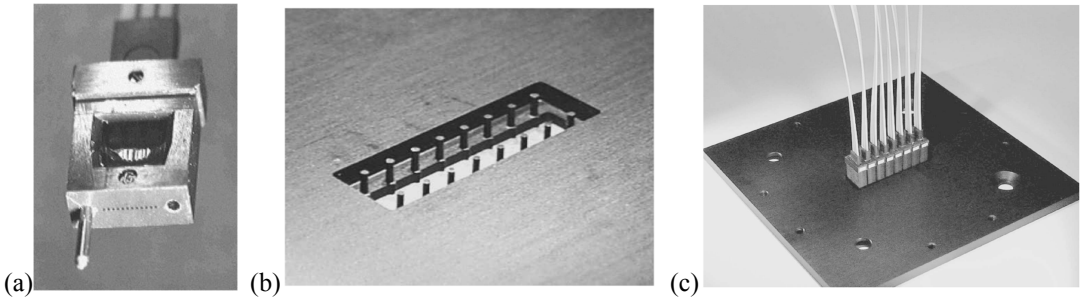

Figure 11: (a) a close-up of an MT connector on a 12-way fiber ribbon.
(b) an interface plate (c) interface plate with 8 fiber ribbons connected

The sum total of the technologies and packaging ideas that have been briefly sketched above is an optoelectronic multi-chip module. The module uses free-space optics to implement data distribution functions and it is equipped with fiber and electronic interfaces. The module is illustrated in Figure 12. In addition to illustrating the ideas already encountered, the figure also makes clear that, in our earliest demonstrator, we expect to package the memory chips and controller on the host PCB rather than mounted on the optical substrate.

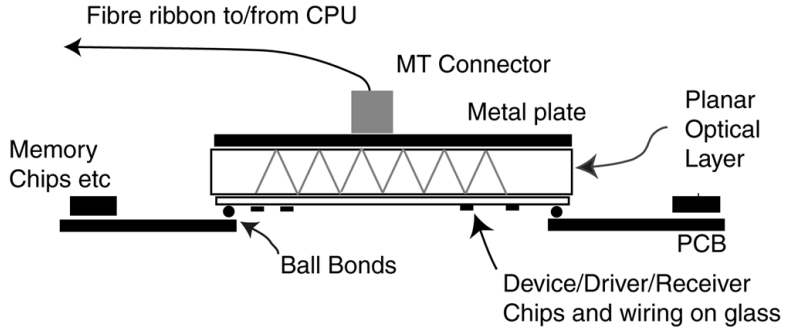

Figure 12: The optoelectronic multi-chip module that forms the heart of the processor-memory interconnection design

6. CONCLUSIONS

We have by no means explored the range of applications of planar optics. These include applications as diverse as neural network and fine-grain computer architectures[15,16] and security and image processing[17]. The potential of free space optics to offer massively parallel interconnections (10^3-10^4 mm^{-2}) has not been explored in this work because we are concentrating on small degrees of fan-out (small numbers of processors) partly for simplicity and partly for reasons of VCSEL power budget. In modulator based systems, such high degrees of parallelism have been demonstrated in planar optics[10,18] and these themselves have intriguing interconnection applications.

This paper has attempted to introduce the planar optical implementation of free-space optics as a candidate for a packageable component in a memory access system. We believe that the functionality of free-space optics will offer significant reduction in latency and that planar optics offers a route to the integration of free-space optics with conventional computer and communications components.

ACKNOWLEDGEMENTS

The authors acknowledge helpful discussions with S. Sinzinger, C. Gimkiewicz and M. Gruber, and are grateful for the technical expertise of D. Hagedorn. This research has been supported by a Marie Curie Fellowship of the European Community Programme *Improving Human Research Potential* under contract number HPMFCT-2000-00700.

REFERENCES

1 D. A. B. Miller and H. M. Özaktas "Limit to the Bit-Rate Capacity of Electrical Interconnects from the Aspect Ratio of the System Architecture", *J. Parallel and Distributed Computing,* **41**, 42-52, 1997
2 J.L. Hennessy and D.A. Patterson, *Computer architecture: a quantitative approach*, Morgan Kaufmann Publ., San Mateo, CA, 1990
3 A. Boxer, "Where buses cannot go", *IEEE Spectrum*, 41-45, Feb. 1995
4 J. Jahns and A. Huang "Planar integration of free-space optical components", *Applied Optics,* **28**, 1602-1605, 1989
5 J. Jahns, "Planar packaging of free space optical components", *Proc. IEEE,* **82**, 1623-1631, 1994
6 P. Lukowicz. "Giga-Cache: A high performance opto-electronic memory architecture", *Proc. Workshop on Critical Technologies in Computing,* SPIE 2000 Annual Meeting, San Diego, CA., 2000.
7 A.W. Lohmann, "Image formation of dilute arrays for optical information processing", *Opt. Commun.*, **90**, 365-370, 1991
8 B. Lunitz and J. Jahns, "Tolerant design of a planar-optical clock distribution system", *Optics Communications,***134**, 281-288, *1997*
9 M. Testorf and J. Jahns, "Paraxial theory of planar integrated systems", *J. Opt. Soc. Am. A*, **14**(7), 1569-1575, 1997
10 S. Sinzinger and J. Jahns, "Integrated microoptical imaging system with high interconnection capacity fabricated in planar optics", *Applied Optics,* **36**(20), 4729-4735, 1997
11 J. Jahns, R.A. Morgan, H.N. Nguyen, J.A. Walker, S.J. Walker, Y.M. Wong "Hybrid integration of surface-emitting microlaser chip and planar optics substrate for interconnection purposes", *IEEE Phot. Techn. Lett.* **4**, 1369-1372 (1992)
12 C. Gimkiewicz, D. Hagedorn, G. Grabosch, J. Jahns, "Heat spreading for vertical-cavity lasers on glass by thermal vias", *IEEE Trans. On Advanced Packaging*, in press, 2001
13 C.Gimkiewicz and D.Hagedorn,"Multichip module with free-space optical interconnects and VCSELs buried into the solder joints ", *Proceedings Micro System Technologies*, Düsseldorf, Germany, 2001
14 P. Lukowicz, S. Sinzinger, K. Dunkel, HD Bauer, "Design of an opto-electronic VLSI/parallel fibre bus" *Journal of Optics A - Pure and Applied Optics* **1**(3), 367-370, 1999
15 M. Gruber, J. Jahns and S. Sinzinger, "Planar integrated optical vector-matrix multiplier" *Appl. Opt*. **39**(29), 5367-5373, 2000
16 D. Fey et al., "Optical Interconnects for Neural and Reconfigurable VLSI Architectures", *Proc. IEEE,* **88***(6),* 838-848, *2000*
17 S. Sinzinger "Microoptical correlators for security applications", OSA Meetings Digest DOMO 2000, 281-282, OSA, Quebec, 2000
18 J. Jahns and B. Acklin, "Integrated planar optical imaging system with high interconnection density", *Optics Letters*, **18**(19), 1594-1596, 1993