

A Three-Dimensional High-Throughput Architecture Using Through-Wafer Optical Interconnect

D. Scott Wills, *Member, IEEE*, W. Stephen Lacy, Christophe Camperi-Ginestet, Brent Buchanan, Huy H. Cat, Scott Wilkinson, Myunghee Lee, Nan Marie Jokerst, and Martin A. Brooke, *Member, IEEE*,

Abstract—This paper presents a three-dimensional, highly parallel, optically interconnected system to process high-throughput stream data such as images. The vertical optical interconnections are realized using integrated optoelectronic devices operating at wavelengths to which silicon is transparent. These through-wafer optical signals are used to vertically optically interconnect stacked silicon circuits. The thin film optoelectronic devices are bonded directly to the stacked layers of silicon circuitry to realize self-contained vertical optical interconnections. Each integrated circuit layer contains analog interface circuitry, namely, detector amplifier and emitter driver circuitry, and digital circuitry for the network and/or processor, all of which are fabricated using a standard silicon integrated circuit foundry. These silicon circuits are post processed to integrate the thin film optoelectronics using standard, low cost, high yield microfabrication techniques. The three-dimensionally integrated architectures described herein are a network and a processor. The network has been designed to meet off-chip I/O using a new offset cube topology coupled with naming and routing schemes. The performance of this network is comparable to that of a three-dimensional mesh. The processing architecture has been defined to minimize overhead for basic parallel operations. The system goal for this research is to develop an integrated processing node for high-throughput, low-memory applications.

I. INTRODUCTION

THE LOGICAL MAPPING for processing elements in most parallel computers is three dimensional. If the implementation of this mapping is constrained to a plane, the processing elements must be separated by longer interconnects. This separation causes increases in signal propagation time between processors (*latencies*), which limits the speed of processor operation. As the number of processing elements is increased, this separation increases, ultimately limiting scalability. In addition, the required wiring density grows with the number of processors. For these reasons, most planar interconnection schemes (printer circuit boards, multichip modules, etc.) are limited in the number of high-speed processing elements that can be implemented.

A three-dimensional interconnection scheme is explored herein using through-silicon wafer optical interconnect. This self-contained system consists of standard silicon foundry circuits which are post-processed with thin film optoelectronic

devices. These optoelectronic devices operate at wavelengths to which the silicon is transparent, thus realizing interconnection through the silicon wafers. This approach allows layers of processing elements and interconnect to be stacked in three dimensions, realizing in physical space the logical mapping. The processes used to integrate these thin film devices onto the silicon circuit are low cost, standard microfabrication techniques.

Network and processing architects are also embracing a step into larger dimensionality through a move from serial processing to parallel processing. Many applications can benefit from large-scale parallelism. But their system requirements vary with a large number of factors, including I/O, storage requirements, and the type of operations required (i.e., integer versus floating point). High-throughput, low-memory problems form one class of applications. This class requires simple parallel tasks to be performed on massive streams of data with less local memory required since operands are provided primarily as I/O. Examples of this application class include image processing (e.g., filtering, edge detection, and convolution), object recognition, and data compression.

This type of problem does not map well onto existing general-purpose parallel computers which are designed for low-throughput, high-memory operation (i.e., a large data set is loaded into memory and extensively processed). Although the I/O bandwidth of these machines is often high, they are inefficient for high-throughput applications. Recent research [1], [9], [12] has explored single-chip processing nodes employing architectures which target more general applications and which use external dense memory chips. But because their access is sequential, dense memory arrays are a poorly utilized silicon resource. Since they are not required for high-throughput applications, the silicon area can be used for additional processing, or eliminated to achieve a lower system cost.

This paper presents a high-throughput processing system which incorporates a new high-bandwidth integrated optoelectronic technology. The use of this technology changes the balance of communication, storage, and processing in the system allowing new approaches to processor and network design. A fine-grain message-passing processing architecture is being designed for handling high message traffic with minimum parallel overhead. Efficient support for task and storage management, communication, and synchronization is

Manuscript received April 25, 1994; revised September 9, 1994.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250 USA.

IEEE Log Number 9411555.

provided. Single cycle task swapping also reduces overhead for short tasks. High-throughput applications and a low-memory processing node make I/O a critical aspect of this architecture. Experiments with example programs indicate that each executed instruction requires approximately 0.1–0.5 words of I/O. If a processor executes at 50 MIPS, the I/O rate can be as high as 25 Words/s. This requires 800 Mb/s of I/O. With dense, three-dimensional arrays of multinode chips, conventional interconnect technology is inadequate.

To satisfy the high I/O requirement, an integrated three-dimensional optical network is being developed, incorporating integrated optical devices developed at Georgia Tech. Arrays of thin film InGaAsP/InP optical emitters and detectors operating at a wavelength of 1.3 μm (to which silicon is transparent) are bonded to silicon chips. These chips are then stacked and staggered to overlap chips in different planes. This forms an offset cube topology similar to a three-dimensional mesh.

This paper describes each aspect of the system design. It begins with a description of through-wafer interconnect and continues with the design of a high-speed analog interface. The physical architecture built using these enabling technologies is then described and the Pica architecture is introduced.

II. THREE-DIMENSIONAL OPTICAL INTERCONNECT

The manufacture of hardware capable of scalable three-dimensional local connectivity has been a challenge for decades. We have demonstrated that through-wafer optical interconnect can be employed to pass information from one layer of silicon to another. This optical interconnect scheme uses thin film InGaAsP/InP emitters and detectors which are bonded directly to each layer in the silicon stack. These emitters and detectors operate at a wavelength of 1.3 μm , to which silicon is transparent, thus enabling vertical optical interconnection of silicon wafers in the three-dimensional stack. The use of thin film emitters and detectors bonded to silicon enables the use of the integration complexity and density of silicon coupled with reliable, low cost fabrication technology for both the circuitry and the integration: standard foundry silicon circuitry and standard microfabrication techniques for the post-processing integration of the thin film devices onto the silicon circuits.

High quality, single crystal thin film InP/InGaAsP materials and devices, either individually or arrays, can be separated from the lattice matched growth substrate using selective etching (called epitaxial lift off [18]), and can subsequently be aligned and bonded onto host substrates such as silicon circuits. This type of mixed-material system is extremely attractive for self-contained three-dimensional silicon systems using through-wafer optical interconnect. Unfortunately, the lattice mismatch and coefficient of thermal expansion mismatch between indium phosphide-based compounds and silicon makes the growth of high-quality indium phosphide-based compounds onto silicon a difficult task. Using epitaxial lift off and thin film bonding, however, which does not suffer from these limitations, we have integrated high quality, single crystal thin film compound semiconductor InGaAsP/InP-based devices onto host substrates which include silicon circuitry, polymers, glass, and lithium niobate.

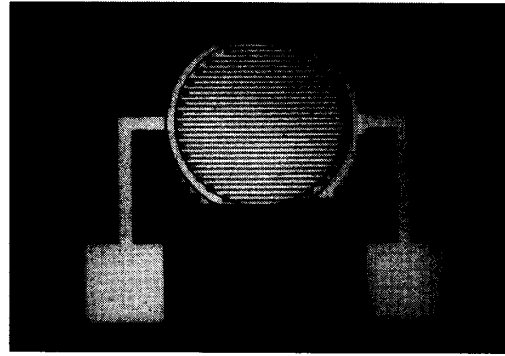


Fig. 1. Thin film InGaAsP MSM detector.

Epitaxial lift off is a promising technique for the integration of compound semiconductors onto smooth substrates since it involves no special growth techniques, has produced high quality GaAs and InP-based materials, and is a relatively simple process with potential for scale up. Epitaxial lift-off utilizes etch selectivity as a function of material composition to separate thin film epitaxial device layers from the growth substrate. In the InP/InGaAsP material system, the etch selectivity between InP and InGaAsP [2] is utilized to form a series of stop-etch layers which are used to separate the thin film epitaxial devices of interest from the growth substrate. These epitaxial device layers of interest generally range from 0.1–5 μm thick. After the epitaxial layers have been separated from the growth substrate, they are commonly handled manually using a thick (100 μm) wax handling layer.

We have developed a polyimide transfer diaphragm that allows the alignment, back contact, and selective deposition of individual devices or arrays of devices [4]. After the epitaxial lift off of mesa etched and surface processed thin film devices, they are attached to a transparent transfer diaphragm. This diaphragm is constructed of either polyimide or mylar, and is supported by a ring of Si on the outer edge. The diaphragm is then inverted (with thin film devices facing the host substrate), and the thin film devices are visually aligned and using a pressure probe, bonded to the host substrate. Current deposition accuracy is to within 1 μm with respect to features on the host substrate.

Thin film InP/InGaAsP/InP p-i-n and InGaAsP MSM detectors which detect at 1.3 μm and p-n homojunction InGaAsP light emitting diodes which emit at 1.3 μm have been fabricated at Georgia Tech. We have also demonstrated through silicon wafer optical communication using these thin film emitters and detectors [3]. These through-wafer demonstrations include a front to back interconnect on a single silicon wafer demonstration as well as a silicon wafer to wafer demonstration. The performance of these thin film devices is comparable to that of devices which have not been separated from the growth substrate [3]. A number of studies have demonstrated that the quality of these thin film materials and devices is not affected by the process of epitaxial lift off and bonding to the host substrate [19]. Fig. 1 is a photomicrograph of an InGaAsP thin film detector which detects at a wavelength of 1.3 μm .

III. ANALOG CIRCUITS

In the Pica architecture described herein, the operand requirement for instruction execution at each node requires 100 Mb/s per channel optical through-wafer communication nodes. To communicate at the approximately 100 Mb/s necessary for each through-wafer optical connection, we have designed a detector amplifier, laser driver, multiplexer/demultiplexer, and timing circuitry with nearly 100 MHz bandwidth (assuming nonreturn to zero (NRZ) encoded serial data transmission) using 0.8 μm CMOS technology. The basic channel architecture proposed is similar to most 100 Mb/s digital fiber optic communication channels with the following notable exceptions. We assume that, due to the tightly controlled environment of the wafer to wafer communication channel, that low Bit Error Rates (BER) will be achieved without the use of error correction codes. We also assume that a global clock signal is available and that only a steady skew will perturb it, i.e., there will be very low clock jitter. The maintenance of low BER requires that signal to noise margins be kept at close to normal CMOS bus levels throughout the channel design procedure. This should be possible with careful detector amplifier design. BER is also impacted by the quality of the clock used to recover the detected digital signals. The presence of only a constant clock skew and no significant jitter will help to maintain low BER since the skew can be canceled by a calibration procedure.

To design the detector amplifiers, a number of assumptions were used to calculate the amplifier parameters. In the through wafer interconnect scheme demonstrated, a 1% optical through wafer efficiency was calculated and indirectly measured. Several factors, such as lensed emitters, use of vertical cavity surface emitting lasers, and use of standard thickness wafers, will improve this figure, but as a worst case we can use it for determining the specifications of the detector amplifier required. Further assumptions of a 50% electrical to optical power conversion in a semiconductor laser emitter, a 5 mW laser power consumption and a 0.5 amp per watt detector efficiency yield a peak detector output current of 12.5 μA . Thus to produce a 5 V signal swing a detector amplifier must have a gain of 400 k Ω . Using the simulated performance of the Hewlett Packard 0.8 μm CMOS foundry process available through the MOSIS service, a three stage open loop amplifier with a gain of 7.4 per stage and total power dissipation of 5 mW will easily meet the gain specifications with a simulated bandwidth of 282 MHz, about three times that needed for a 100 Mb/s NRZ encoded serial signal. Similar amplifiers fabricated by the authors have had equivalent input noise currents of 10–100 nA. This gives a signal to noise ratio of 100–1000 or 40–60 dB which is superior to the noise margin typical for a CMOS data bus.

The primary concern with clock distribution is the calibration of the skewed clocks. We can assume that each chip in the processor array will have a slightly different but essentially constant clock skew. Adjacent processors will have very similar skews, and careful design of the clock distribution network will ensure that the skew between locally connected processors on any one processing plane will be negligible.

Thus each processor must correct the skew in the clock from processors on the plane above or below it in a stack, but will need only one correction for all four connected processors on the same plane. One high end solution to this incorporates a single analog phased locked loop onto each chip in the array, and corrects the skew with a 4 or 5 bit digital to analog converter (DAC) to set the skew. This scheme assumes a short synchronization period at power-up, and a supply voltage stable to one least significant bit (LSB) for the DAC.

None of this proposed scheme presents a technological challenge, and simplified schemes may well suffice. The scheme will operate unless the skew between adjacent chips becomes significant; at that point one more correction phase locked loops per channel will be needed. However, that problem will not occur until clock speeds show significant skew in less than a centimeter, which would require GHz clocks.

The multiplexing of 100 Mb/s NRZ encoded data back to the 50 MHz chip clock data is simplified by using the same clock speed for the MUX/DEMUX as the rest of the processor. This clock speed is about one half of the 100 MHz 31 stage ring oscillator frequency of the HP 0.8 μm CMOS process.

The design of laser drivers is similar to digital pad drivers. We propose to add a variable current ECL like differential output stage to allow a variable laser drive current. This will allow the power dissipation of the lasers to be controlled and adjusted to minimize heat dissipation. It will require a special low voltage, high current dc power bus of around 2 V for the lasers. Fortunately, the 2.5 mA of average laser drive current required for this application is well within the capabilities of CMOS drivers.

IV. NETWORK ARCHITECTURE

In order to meet the I/O requirements of several nodes per chip, a high-bandwidth, three-dimensional network is required. Traditional wire interconnects which are limited to perimeter chip contacts cannot scale as transistor densities increase. New MCM technologies (e.g., bump bonding) are not suitable for scalable three-dimensional networks. Existing optical interconnect techniques such as index-guided and free-space networks are difficult to incorporate in three-dimensional interconnection networks. We are currently developing a three-dimensional optoelectrical network using hybrid integrated optoelectronic devices and through-wafer transmissions.

A. Background

To build a system, processing elements are designed and fabricated using existing Si VLSI techniques. No bonding pads are included except for power contacts. (Probe pads are provided for testing.) Instead, holes in the overglass allow GaAs and InP-based emitters or detectors to be attached in a post-processing step. Driver circuits, included on the silicon, interface parallel buses from routers on the chip to the faster optical devices (0.1–1.0 Gb/s).

After the Si chips containing the optical devices are fabricated and tested, they are attached to a larger silicon substrate. This provides physical support and also participates in power

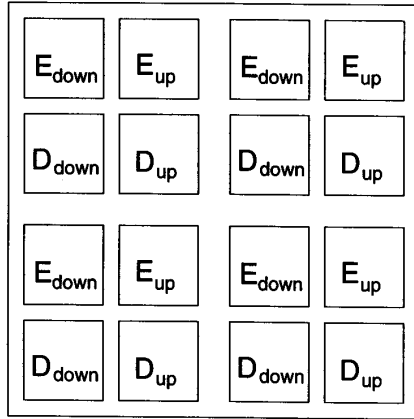


Fig. 2. Chip device pattern.

distribution and cooling. A fully populated substrate forms one plane of the system. Completed substrates are then stacked to create the third dimension. To facilitate manufacturing, all chips and substrates are identical.

Each chip is broken into four quadrants which overlap with eight neighboring chips (four in the plane above, four in the plane below). Although any number of emitter and detectors can be included in each quadrant, the minimum case requires two emitter/detector pairs, shown in Fig. 2. In order for emitter/detector pairs to be correctly aligned, the chips must be offset. To achieve this using identical plane substrates, a spacing equal to one half the chip width plus inter-chip spacing is added to two nonopposing sides of the substrate. Then alternating planes are rotated 180 degrees during assembly providing the correct device alignment.

System I/O is provided by I/O layers which inject data directly into the network on the top and bottom surfaces of the stack. These layers can interface to high-bandwidth electrical or optical connections. Alternatively, the input surface can be covered by visible light detectors for direct focal-plane imaging. The output surface can incorporate a frame buffer for direct video output. The sides of the stack are used for power distribution and cooling connections. Due to the dense packaging, liquid cooling between layers is anticipated.

B. Offset-Cube Topology

Since each chip in the system contains several nodes, the communication network is divided into two parts. The inter-chip network delivers messages between the chips containing the source and destination nodes. The intra-chip network is used to communicate messages within the source and destination chips. Messages sent between nodes within a chip are routed entirely via the intra-chip network.

Communication within a chip is accomplished using a wire-only network. This type of network is well-studied [7], [13]. Since on-chip wire density is high and the number of nodes per chip is initially low, the topology of the intra-chip network is straightforward. Each chip connects to eight neighboring chips: four in the plane above and four in the

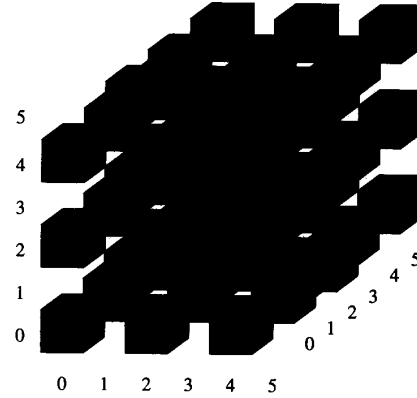


Fig. 3. A 3-ary offset cube topology.

plane below. There are no inter-chip connections within a substrate plane. This interconnection forms an offset cube topology which bears similarities to the well-studied k -ary 3-cube. While the physical architecture of the inter-chip network is not symmetric in the vertical and horizontal dimensions, the offset cube topology implemented by this network is isotropic.

The topology is formally defined as follows. Consider a set of k^2L vertices (representing processing elements) where k is the length of each side of a square vertex array on each layer and L is the number of layers. The vertices are named by triples $\langle x, y, l \rangle$ where l is the layer number between 0 and $(L - 1)$ inclusive. For even values of l , $x = 2i + 1$ and $y = 2j + 1$ where i and j are the vertex's coordinates within a layer assuming values between 0 and $k - 1$. For odd values of l , $x = 2i$ and $y = 2j$ for i and j between 0 and $k - 1$. The offset cube topology is defined by the edges connecting each vertex $\langle x, y, l \rangle$ with its neighbors $\langle x \pm 1, y \pm 1, l \pm 1 \rangle$. Boundary vertices have less than eight edges, and are defined in six overlapping sets.

- 1) vertices $\langle 0, y, l \rangle$ are connected with its neighbors $\langle 1, y \pm 1, l \pm 1 \rangle$
- 2) vertices $\langle 2k - 1, y, l \rangle$ are connected with its neighbors $\langle 2k - 2, y \pm 1, l \pm 1 \rangle$
- 3) vertices $\langle x, 0, l \rangle$ are connected with its neighbors $\langle x \pm 1, 0, l \pm 1 \rangle$
- 4) vertices $\langle x, 2k - 1, l \rangle$ are connected with its neighbors $\langle x \pm 1, 2k - 2, l \pm 1 \rangle$
- 5) vertices $\langle x, y, 0 \rangle$ are connected with its neighbors $\langle x \pm 1, y \pm 1, 1 \rangle$
- 6) vertices $\langle x, y, L - 1 \rangle$ are connected with its neighbors $\langle x \pm 1, y \pm 1, L - 2 \rangle$.

For a symmetric cube, $L = 2k - 1$ creating a k -ary offset cube. The offset cube topology is isotropic, and is illustrated in Fig. 3. In a k -ary 3-cube, a shortest path routing algorithm uses channel routes to reduce the difference in coordinates of the source and destination vertices. A well-performing deterministic shortest path algorithm reduces the coordinate differences one at a time (i.e., x , then y , then z , or *dimension order routing*).



Fig. 4. Routing paths in an offset cube.

A shortest path routing algorithm for an offset cube is more complicated since each channel hop changes all three coordinates. If one coordinate difference reaches zero, it must be increased in the next hop to reduce differences in other dimensions. This creates “bouncing” routes where a message alternates between neighboring planes as it moves through one dimension. Several routing paths are shown in Fig. 4.

The minimum path cost of routing a message between vertices A and B on an offset cube is:

$$\text{hops}_{\min} = \max(|X_b - X_a|, |Y_b - Y_a|, |Z_b - Z_a|).$$

Deadlock avoidance in an offset-cube is similar to that in k -ary 3-cubes. Existing techniques [8] are being adapted to this network.

C. Network Performance

The analysis of this network is similar to that of wire-only networks [7], except that the optical channels require a different set of assumptions.

- 1) Unlike the wire density limits for wire networks, optical channels are limited entirely by characteristics of the transmitter and receiver. The communication medium (i.e., the space through which light travels) does not enforce any fundamental density or bandwidth limits.
- 2) The channel delay is not determined by channel length. Since the messages travels fast (near c), and the channel length is small (around 1 mm), the time of flight in a channel is around three picoseconds. The channel delay is dominated by transmitter and receiver response time (around 1 nS).
- 3) Power and heat dissipation are the dominant issues in communication costs. GaAs LED's have a low quantum efficiency (<3%). GaAs lasers are more efficient (75%), but still dissipate several milliwatts during operation. The cumulative effect of thousands of lasers operating in a few hundred cm^{-3} will require special consideration.

Performance of the offset cube topology has been analyzed through simulation. This section presents a summary of the results of this analysis. The complete details are available in [10].

To compare the k -ary 3-cube and offset cube topologies, a simulator has been constructed which analyzes network performance under similar conditions and loads. In these experiments, a 16-ary 3-cube is compared with a 13-ary offset cube. A random traffic load is simulated. Wormhole routing is employed with eight virtual channels for each physical channel. All message lengths are 25 flits.

Fig. 5 summary-plot compares the performance of deterministic and adaptive routing algorithms on the two topologies. For the deterministic case, the 16-ary 3-cube uses dimension-order routing. The offset cube routes along a diagonal shortest path. The offset cube utilizes less than half of the capacity of the 3-cube.

The capacity loss is due to loading in the center of the network. For random traffic, deterministic routing on a 3-cube does a reasonable job of distributing traffic across the entire network. The offset cube's diagonal routing creates much higher traffic in the center of the network. This has been verified by comparing cross-sectional loading profiles for a 3-cube with dimension-order and diagonal routing algorithms.

To better utilize the capacity of an offset cube network, a simple adaptive routing algorithm is employed. Local information (the number of virtual channels in use) is used to select between shortest paths at each vertex. The adaptive offset cube performance exceeds that of the best 3-cube strategy. The adaptive 3-cube performance is lower than the deterministic case because of increased diagonal routing.

V. PROCESSOR ARCHITECTURE

The Pica execution architecture is designed for handling high message traffic consisting of small, ephemeral tasks. In order to achieve acceptable efficiency in this fine-grain domain, parallel overhead must be reduced to the minimum achievable level. Complex mechanisms to support general purpose applications are replaced by simpler, lower cost mechanisms for high-throughput problems.

The Pica execution architecture is designed specifically for high-throughput, low-memory operation. The design of a Pica node begins with a minimal sequential core architecture. Pica provides low overhead support for communication, synchronization, naming, and task and storage management. A small amount of memory (4096 36-b words) and a network interface/router complete the node. This node complexity can be implemented using a fraction of the transistors available on a chip in current technology. This allows multinode chips; the prototype chip will contain four nodes.

The Pica architecture is designed to form a dense, three-dimensional computational array for processing high-throughput data streams. While less general than other MIMD architectures, it is more efficient for this application area. The execution model supported by Pica is more flexible than other high-throughput architectures (e.g., systolic arrays, static dataflow).

The basic functional blocks of the Pica microarchitecture are shown in Fig. 6. The network router routes messages through the node, forming that node's contribution to the communication network. The router implements a simple adaptive routing strategy based on current local virtual-channel allocation. The network interface buffers incoming messages and signals the context manager that a context is required. When it obtains access to local memory, the network interface writes the message contents directly into the allocated, fixed-length context. The datapath consists of a 32-b integer ALU and shifter, and special-purpose registers. Operands are accessed

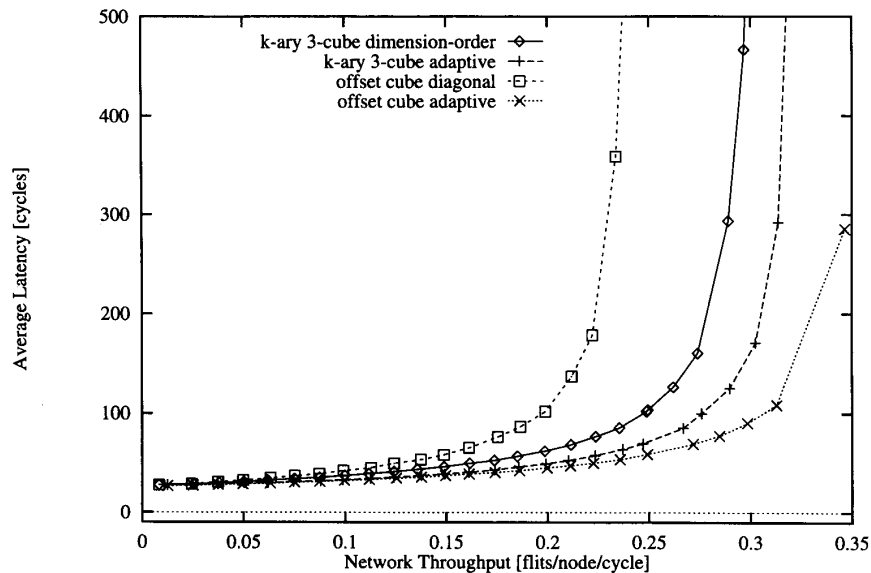


Fig. 5. Performance of offset cube and 3-cube.

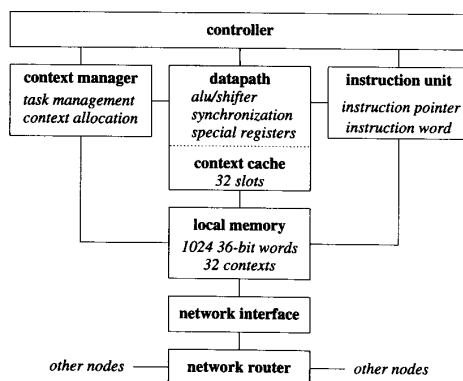


Fig. 6. Pica microarchitecture.

from a 32 word context cache, which supports two read and one write accesses on each cycle. The instruction unit fetches and decodes instructions for execution. In order to keep design complexity and task swapping overhead low, the datapath implementation is not pipelined. The context manager serves three functions: 1) it maintains a queue of suspended and ready tasks for execution, 2) it allocates task storage for incoming messages and deallocates storage as the tasks complete, and 3) it arbitrates requests by both the network interface and cache controller for control of the local memory bus.

VI. STATUS AND SUMMARY

The communications and processing components of a highly parallel, optically interconnected Pica architecture have been presented. The optical communication has been demonstrated using through-wafer thin film optoelectronic devices which operate at a wavelength to which silicon is transparent. Analog circuits have been designed which operate at data rates of 100 Mb/s which interface to the Pica network interface. The initial

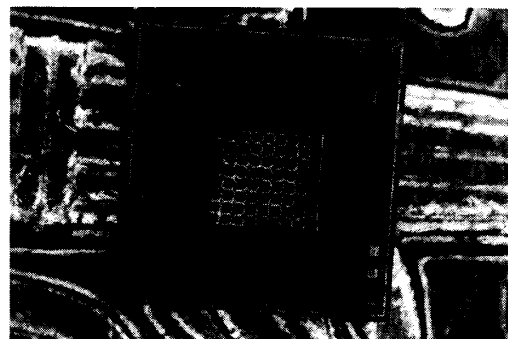


Fig. 7. Eight by eight array of GaAs-based p-i-n directly on top of silicon amplifier circuits.

processor design under implementation targets 73 Kb of static storage (2048 words, 442K transistors) and 100K transistors for the datapath and controller. This should allow a four-node chip using current technology. Test components of the processing node and optical network are being fabricated. Fig. 7 is a photomicrograph of an 8×8 array of GaAs-based p-i-n detectors integrated directly on top of a silicon amplifier circuitry.

Simulations of the network performance are being conducted [16], [17], [11]. Plans for a full-scale prototype are underway, but are still preliminary. A full-scale prototype system will have the following characteristics. The node being designed contains 2048 36-bit words of local memory (256 contexts), a 32-bit integer processor, and a network interface. The estimated instruction rate is 50 MIPS. A chip contains four nodes and 3.2 Gb/s optical I/O bandwidth. A processing plane contains 64 chips (256 nodes, 12 800 MIPS) and measures approximately 10 by 10 cm. 16 planes contain 1024 chips (4096 nodes, 204 800 MIPS) and fit inside a cube 10 cm on a side. 819.6 Gb/s of I/O bandwidth is

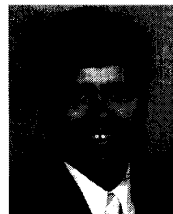
available from chips on the top and bottom of the cube. Sides of the cube are used for power and cooling mechanical connections. The object recognition application domain has been selected for evaluating the Pica. This domain is chosen because a) massively parallel algorithms are available, b) high-throughput processing is required, and c) extensive floating-point performance is not needed. An assembler and instruction-level simulator are being used to develop application programs and evaluate the architecture [15]. A compiler for a higher-level language (Pica-C) is in progress.

Using these tools, several important image processing operations and applications have been implemented, including convolution, relaxation, FFT, edge detection, JPEG compression, and array manipulation [14]. The largest implemented application for Pica is a Maximum Likelihood Expectation Maximization algorithm for positron emission tomography [5], developed in conjunction with the author's research group and Crawford Long Hospital in Atlanta. This application, which reconstructs 21 slices of 256×256 pixels over 120 angles, offers an effective speedup of 900 as compared with the sequential workstation currently in use.

Using the instruction-level simulator, actual message traces from these applications can be collected and used by a second generation network simulator that is nearing completion. This network workload provides a more realistic comparison of offset-cube networks with other topologies (e.g., k -ary, 3-cubes).

REFERENCES

- [1] W. C. Athas and C. L. Seitz, "Multicomputers: Message-passing concurrent computers," *IEEE Computer*, vol. 21, no. 8, pp. 9-24, Aug. 1988.
- [2] G. Augustine, N. M. Jokerst, and A. Rohatgi, "Single crystal thin film InP: Fabrication and absorption measurements," *Appl. Phys. Lett.*, vol. 61, no. 12, pp. 21-23, Sept. 1992.
- [3] K. H. Calhoun, C. B. Camperi-Ginestet, and N. M. Jokerst, "Vertical optical communication through stacked silicon wafers using hybrid monolithic thin film InGaAsP emitters and detectors," *IEEE Photon. Technol. Lett.*, vol. 5, no. 2, pp. 254-257, Feb. 1993.
- [4] C. Camperi-Ginestet, M. Hargis, N. Jokerst, and M. Allen, "Alignable epitaxial liftoff of GaAs material with selective deposition using polyimide diaphragms," *IEEE Photon. Technol. Lett.*, vol. 3, no. 12, pp. 1123-1126, Dec. 1991.
- [5] C. M. Chen, S. Y. Lee, and Z. H. Cho, "Parallelization of the EM algorithm for 3-D PET image reconstruction," *IEEE Trans. Med. Imaging*, vol. 10, no. 4, pp. 513-521, Dec. 1991.
- [6] J. L. Cruz-Rivera, E. V. R. DiBella, D. S. Wills, T. K. Gaylord, and E. N. Glytsis, "Parallelized implementation of the maximum likelihood expectation maximization algorithm on a fine-grained optically interconnected architecture," submitted to *IEEE Trans. Med. Imaging*, 1994.
- [7] W. J. Dally, "Performance analysis of k -ary n -cube interconnection networks," *IEEE Trans. Comput.*, vol. 39, no. 6, pp. 775-785, June 1990.
- [8] W. J. Dally and C. L. Seitz, "Deadlock-free message routing in multiprocessor interconnected networks," *IEEE Trans. Comput.*, vol. C-36, no. 5, pp. 547-553, May 1987.
- [9] W. J. Dally, J. A. S. Fiske, J. S. Keen, R. A. Lethin, M. D. Noakes, P. R. Nuth, R. E. Davison, and G. A. Fyler, "The message-driven processor: A multicomputer processing node with efficient mechanisms," *IEEE Micro*, vol. 12, no. 2, pp. 23-39, Apr. 1992.
- [10] M. Grossglauser and D. S. Wills, "Performance analysis of an offset cube network using optical interconnect," VLSI Architectures Group Technical Report, Jan. 1993.
- [11] W. S. Lacy, J. Cruz-Rivera, and D. S. Wills, "A scalable optical interconnection network for fine-grain parallel architectures," submitted to *IEEE Trans. Comput.*, Mar. 1994.
- [12] D. May, R. Shepherd, and P. Thompson, "The T9000 transputer," in *Proc. 1992 IEEE Int. Conf. Computer Design: VLSI in Computers and Processors*, Oct. 1992, pp. 209-212.
- [13] D. A. Reed and R. M. Fujimoto, *Multicomputer Networks: Message-Based Parallel Processing*, Scientific Computation Series. Cambridge, MA: MIT Press, 1987.
- [14] G. K. Wallace, "Overview of the JPEG (ISO/CCITT) still image compression standard," in *Current Overviews in Optical Science and Engineering I*, R. Feinberg, Ed. Bellingham, WA: SPIE Optical Engineering Press, 1990, pp. 358-371.
- [15] D. S. Wills, H. Cat, J. Cruz-Rivera, W. S. Lacy, J. M. Baker, J. Eble, A. Lopez-Lagunas, and M. Hopper, "The pica architecture and its application," submitted to *IEEE Trans. Parallel Distributed Syst.*, Feb. 1994.
- [16] D. S. Wills and M. Grossglauser, "A scalable optical interconnection network for fine-grain parallel architectures," in *Proc. 1993 Int. Conf. Parallel Processing*, Aug. 16-20, 1993, pp. 1-154-1-157.
- [17] ———, "A three-dimensional optical interconnection network for fine-grain parallel architectures," in *Proc. IEEE Lasers and Electro-Optics Summer Topical Meeting on Hybrid Optoelectric Integration and Packaging*, July 26-28, 1993, pp. 21-22.
- [18] E. Yablonovitch, T. J. Gmitter, J. P. Harbison, and R. Bhat, "Extreme selectivity in the liftoff of epitaxial GaAs films," *Appl. Phys.*, vol. 51, no. 26, pp. 2222-2224, 1987.
- [19] E. Yablonovitch, E. Kapon, T. Gmitter, C. Yun, and R. Bhat, "Double heterostructure GaAs/AlGaAs thin film diode lasers on glass substrates," *IEEE Photon. Technol. Lett.*, vol. 1, no. 2, pp. 41-42, Feb. 1989.



D. Scott Wills (S'80-M'90) received the B.S. degree in physics from Georgia Institute of Technology in 1983. He received the S.M. (1985), E.E. (1987), and Sc.D. (1990) degrees in electrical engineering and computer science from Massachusetts Institute of Technology.

He is an Assistant Professor of Electrical and Computer Engineering at Georgia Institute of Technology. His research interests include parallel architectures, multicomputer networks, VLSI design, and parallel compilers.

Dr. Wills is a member of the IEEE Computer Society. He is also chairman of the Atlanta Chapter of IEEE Computer Society.

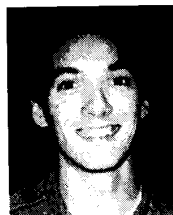


W. Stephen Lacy received the B.S. degree in electrical engineering from Christian Brothers University, Memphis, TN, in 1991, the M.S.E.E. degree in electrical engineering from the Georgia Institute of Technology in 1993, and is currently pursuing the Ph.D. degree in electrical engineering with a minor in computer science at Georgia Tech.

His research interests include circuits and packaging for high-performance VLSI, multicomputer interconnection networks, trace-driven network simulation, and architectures for fine-grained parallel

processing.

Mr. Lacy is a member of the Tau Beta Pi, Alpha Chi, and Phi Beta Kappa Academic Honor Societies.

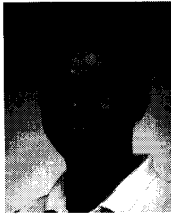


Christophe Camperi-Ginestet received the Engineer Diploma in physics of semiconductor devices from the Institut National des Sciences Appliquées (INSA), Toulouse, France, in 1990 and the Ph.D. degree in electrical engineering from the Institut National des Sciences Appliquées (INSA), Toulouse, France, in 1993.

He is currently Research Engineer at the Georgia Institute of Technology. His research interests include optimization of thin films devices for three-dimensional interconnects, integration of optoelectronics devices onto processing circuits (GaAs, Si).

Brent Buchanan received the B.S. degree in applied physics (June 1988, High Honors), B of Electrical Engineering (June 1991, Highest Honors), and M.S. degree in electrical engineering (December 1991) from the Georgia Institute of Technology and is currently pursuing the Ph.D. degree in electrical and computer engineering at Georgia Tech.

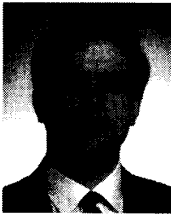
His research interests include VLSI hardware for low-level machine vision and analog pre-processing circuits for Digital Signal Processors.



Huy H. Cat received the B.S. degree in computer engineering from University of Central Florida in 1992, the M.S.E.E. degree in electrical engineering from the Georgia Institute of Technology in 1994, and is currently pursuing the Ph.D. degree in electrical engineering.

His research interests include architectures for fine-grained parallel processing, high performance VLSI architectures, computer modeling.

Mr. Cat is a member of the Tau Beta Pi.



Scott Wilkinson received the Bachelor of Electrical Engineering from Georgia Tech in 1990.

He is a doctoral candidate in electrical engineering at the Georgia Institute of Technology. His area of research is in integrated optoelectronics.



Myunghee Lee received the B.S. degree in electronics from Hanyang University, Seoul, Korea, in 1984 and the M.S.E.E. degree from Arizona State University, Tempe, AZ, in 1990.

He is a Ph.D. candidate of electrical and computer engineering at the Georgia Institute of Technology. He worked at Digital Equipment Korea as a digital circuit design engineer about 5 years before he started the M.S. degree. He also worked at IBM T. J. Watson Research Center as an Engineer during 1990-1991. Since he joined Georgia Tech in 1991,

his work has focused on a high-speed optical receiver for SONET and FDDI. He is also involved in various other projects.



Nan Marie Jokerst received the B.S. degree in physics and mathematics from Creighton University in 1982, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California in 1984 and 1989, respectively.

She is an Associate Professor of Electrical and Computer Engineering at the Georgia Institute of Technology and is associated with the Manufacturing Research Center and the Microelectronics Research Center at Georgia Tech. She was named a National Science Foundation Presidential Young

Investigator in 1990, a DuPont Young Professor in 1989, a Newport Research Award winner in 1986, and a Hewlett Packard Fellow in 1983. She has also been recognized for her teaching accomplishments, which include second place in the Eta Kappa Nu C. Holmes MacDonal Young Professor in Electrical Engineering National Competition in 1992 and 1993 and the award for Best Teacher in Electrical Engineering at Georgia Tech in 1990. Her work has focused on fabrication, integration and application of thin film semiconductor materials and devices. Her work in this area includes manufacturing technology for electronic and optoelectronic integration and solderless interconnect, and three-dimensional integration of both electrical and optical signal layers.



Martin A. Brooke (S'85-M'87) received the B.E.(Elect.) degree (1st. Class Hons.) from Auckland University, New Zealand, in 1981. He received the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California in 1984 and 1988, respectively. His doctoral research focused on reconfigurable analog and digital integrated circuit design.

He is currently Associate Professor of Electrical Engineering at the Georgia Institute of Technology.

He won the 1990 National Science Foundation Research Initiation Award, and the 1992 IEEE Midwest Symposium on Circuits and Systems, Myril B. Reed Best Paper Award. He has 3 U.S. patents awarded, and 1 pending. He has published 32 refereed papers in journals and conference proceedings, and articles on his work have appeared in 15 different trade publications.

Dr. Brooke is a member of the INNS.