Advanced Ultra High Speed Processor Technologies

Final Report

Contract No. N00019-96-C-0032
Contract Amount: $70,000

Period of Performance: 12-18-95 to 06-18-96

Sponsor:

Naval Surface Warfare Center
10901 New Hampshire Avenue
Silver Spring, MD 20903

Technical Monitor:

Mark Silbert
(215) 441-2556

Contractor:

Physical Optics Corporation
Engineering and Products Division
20600 Gramercy Place, Bldg. 100
Torrance, CA 90501-1821

Principal Investigator:

Andrew Kostrzewski, Ph.D.
(310) 320-3088

June 18, 1996

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	06/18/96	Final, 12/18/95 to 06/18/96

4. TITLE AND SUBTITLE
Advanced Ultra High Speed Processor Technologies

5. FUNDING NUMBERS
N00019-96-C-0032

6. AUTHOR(S)
Andrew Kostrzewski, Ph.D.
Jeongdal Kim, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Physical Optics Corporation
Engineering and Products Division
20600 Gramercy Place, Bldg. 100
Torrance, California 90505

8. PERFORMING ORGANIZATION REPORT NUMBER
3363

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Surface Warfare Center
10901 New Hampshire Avenue
Silver Spring, MD 20903

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In Phase I of this project, POC designed, fabricated, and tested optical interconnects, and designed a graphics processor based on field-programmable gate arrays (FPGAs). POC's board-to-board connection approach is based on multi-channel integrated optical waveguides with novel optoelectronic active connectors, allowing multiple simultaneous data transfers among many boards. Data is transferred from chip to chip through optical data channels within an integrated chip. In the electronic processing system design, POC developed a preliminary design of a multiprocessor system suitable for single instruction multiple data (SIMD) and multiple instruction multiple data (MIMD) with video random access memory (VRAM) and static random access memory (SRAM) for main memory, look-up table processing, and display interface.

14. SUBJECT TERMS
Field-Programmable Gate Array (FPGA), SIMD, MIMD, VRAM, SRAM

15. NUMBER OF PAGES
37

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

TABLE OF CONTENTS

1.0 INTRODUCTION

The need for higher processing power has led to the development of new integrated circuits for real-time processing of video/graphic information, and introduction of high speed rendering accelerators for 3-D graphics has moved the PC market in a new direction. However, more demanding applications are emerging that require even higher processing speeds, which can be achieved only by combining more processing resources into a compact system. A multiprocessor parallel system with both local and shared memory is therefore essential to the future development of high-end processing systems. The processing speed achievable by using many ICs increases rapidly, but a major roadblock is the limitation of the electrical interconnections. Because electrical system-to-system, board-to-board, and chip-to-chip interconnections are slowed by low bandwidth, cross talk and channel density limitations, next generation processing systems must incorporate optical interconnects, especially for digital signal transmission over greater interconnection distances and higher bandwidths. Fiber optic data transmission has revolutionized the telecommunications industry. Now we need optical technology to do the same on the local board-to-board and even chip-to-chip levels.

Recent years have seen progress toward the use of optical links over shorter and shorter distances. The undisputed superiority of optical transmission in long-haul telecommunication is expanding now into shorter distances of 1-to-100 meters.

In this project, POC has concentrated on developing even shorter range optical data communications for board-to-board (<1 m distance) or even shorter chip-to-chip or multi-chip module (MCM)-to-MCM (<10 cm).

Optical interconnections can be implemented for board-to-board interconnects as well as for chip-to-chip interconnects. The second is more demanding because of the necessity to integrate electronic components with optoelectronic data transmission drivers. Initially, such development is possible using MCM technology, but the ultimate goal is to build ICs with "optical pins" for very high bandwidth data transfer directly on the IC. Such hybrid ICs will require development of special printed circuit boards to support optical data transmission by means of optical fibers or optical waveguides. The major benefits of the hybrid ICs will accrue from their very high data bandwidth transfer and from the significant reduction of real estate achieved by time/wavelength multiplexing of many signals on a single optical channel.

High speed graphics application can also benefit from parallel processing architectures with optical interconnects. The most powerful paradigm of this scheme is multiple instruction multiple data (MIMD). Recently, *massively parallel processing* (MPP) systems have adopted scalable and latency-tolerant architectures using VLSI silicon, GaAs technologies, high-density packaging, and optical interconnection technologies. MPP computers are expected to achieve teraFLOPS (10^{12} floating-point operations per second) rates within a year or so, enhancing graphics handling capabilities far beyond what has been possible. Thus, any chip or board level innovation must take into account system level considerations of parallel architecture.

In Phase I, a hybrid processing system combining high speed electronic processing with innovative high speed optical interconnections has been designed, along with a parallel processing system implemented with field programmable gate arrays (FPGAs). Optical interconnects are used at the board-to-board and chip-to-chip levels. Optical interconnects at chip-to-chip levels require development of an "optical pin" equivalent to a large number of electronic pins (e.g., an entire 32-bit data bus can be multiplexed on a single optical data channel). Optical interconnects are also being applied for board-to-board communication. High optical data transfer bandwidth and high channel density allow implementation of nonblocking interconnection topologies such as large

scale crossbars that significantly improve multiprocessor connectivity in comparison to hypercube and mesh interconnection topologies.

2.0 PHASE I RESULTS

This section presents the results of the Phase I work.

2.1 Highlights of Phase I Results

Developments in this six month project were in response to two main tasks:

- Develop a hybrid optoelectronic approach to improve the performance of electronic processing systems
- Design a parallel processing system to implement this optoelectronic approach.

The two major accomplishments in Phase I were:

- POC designed, fabricated, and tested optical interconnects as discussed in Sections 2.2 through 2.4.
- POC designed an MIMD graphics processor based on field-programmable gate arrays as discussed in Section 2.9.

POC's board-to-board connection approach is based on multi-channel integrated optical waveguides with novel optoelectronic active connectors, allowing multiple simultaneous data transfers among many boards. Data is transferred chip-to-chip through optical data channels within an IC. One optical data channel can be used to connect a large number of equivalent electronic channels by means of time or wavelength division multiplexing. Initially, optical chip-to-chip data transfer will take advantage of MCM technology, in which an electronic IC is integrated with parallel-to-serial converters and an optoelectronic driver with a laser diode for data transmission. On the receiver side, we will integrate a photodiode with an amplifier and signal conditioning/recovery circuitry and a serial-to-parallel converter. The design of the optical data transfer subsystem to be used at the board-to-board and chip-to-chip levels is outlined in Section 2.2.

In the electronic processing system design, we concentrated on 3-D graphic processors, with the emphasis on the hardware architectures. As a result of this study, we have developed a preliminary design of a multiprocessor system suitable for single instruction multiple data (SIMD) and multiple instruction multiple data (MIMD) configurations. Each processing node will include a field programmable gate array (FPGA) with video random access memory (VRAM) and static random access memory (SRAM) for main memory, look-up table processing, and display interface (see Sections 2.7 and 2.8).

The following discussion first outlines the hybrid optoelectronic processing system, followed by the design of SIMD and MIMD multiprocessor systems.

2.2 General Overview of Optical Interconnect Design

Purely electronic interconnects are limited by their long propagation times, low bandwidths, large clock skews, and large resistor-capacitor (RC) time constants. These limitations cause serious bottlenecks even in the most advanced electronic board-to-board interconnects. In this project,

POC has developed an optical communication system with data speeds well beyond those of existing electronic systems. In the initial study stage of this project, POC has selected a hybrid solution for the design of the high speed processor.

The multiprocessor system combines electronic processing hardware with high-bandwidth optical interconnects. This combination permits very high system throughput by removing most of the limitations that characterize electronic interconnects (see Figure 2-1).

The preliminary design was a multi-board system as illustrated in Figure 2-1. Board-to-board communication is through high speed optical data and address buses. Slower control channels use electronic connections. Each board contains about four processing elements (PEs), which exchange data at high speed over dedicated optical lines as shown in Figure 2-2. Both direct modulation and external modulation techniques are used. Chip-to-chip communication employs external modulation, since it is easier to integrate external modulators with VLSI structures. Board-to-board interconnects use direct modulation because of their higher power requirements (which are due to fan-out).

The optical interconnects are based on integrated optics to ensure rugged connectorization to hold the laser diodes (LDs) and photodetectors (PDs) in alignment. Another solution is to fix the LDs and PDs in place on the board to maintain higher communication speed. POC can achieve high density optical waveguide packaging (100/mm). The processing bands are connectorized for easy insertion and removal.

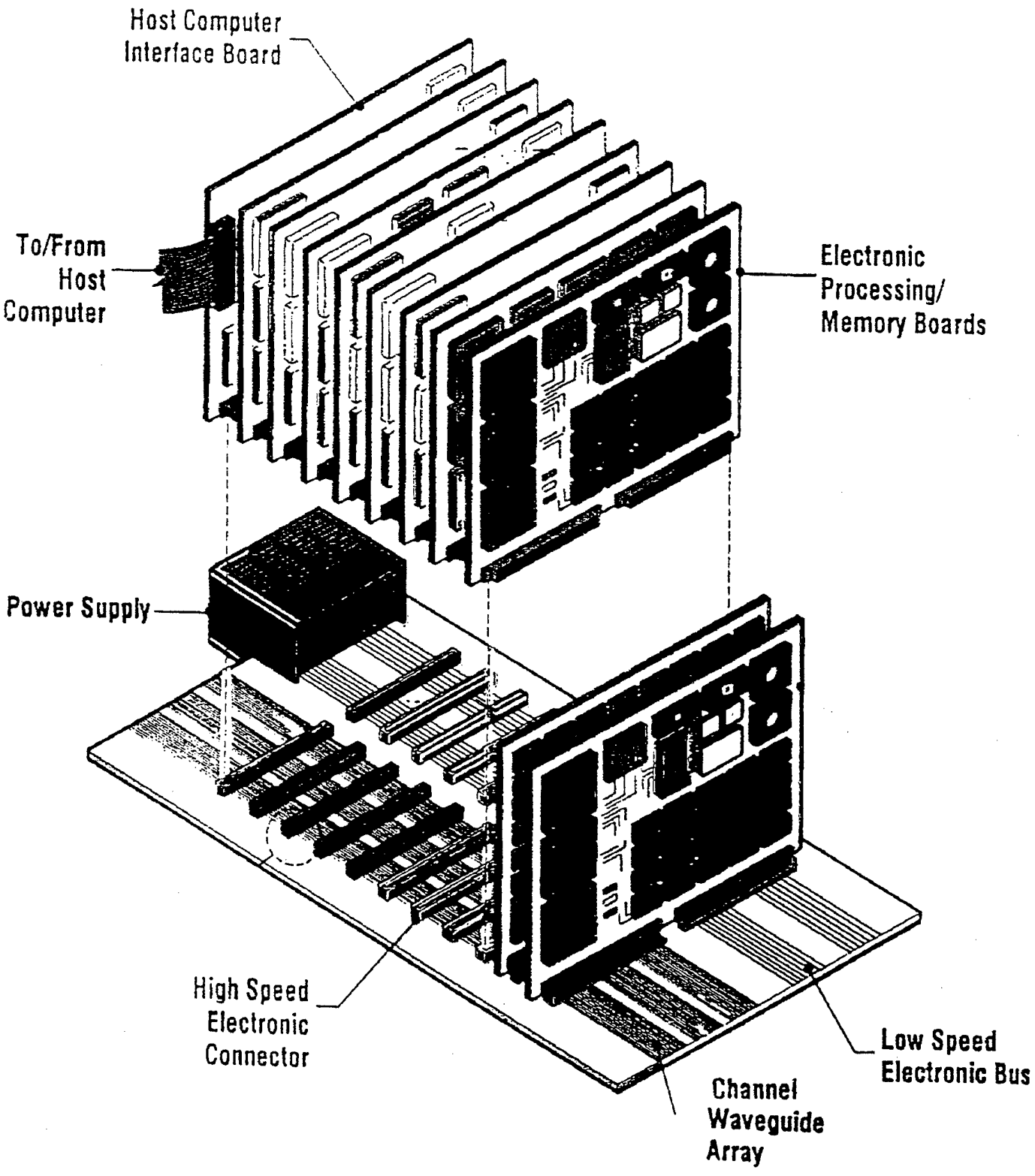
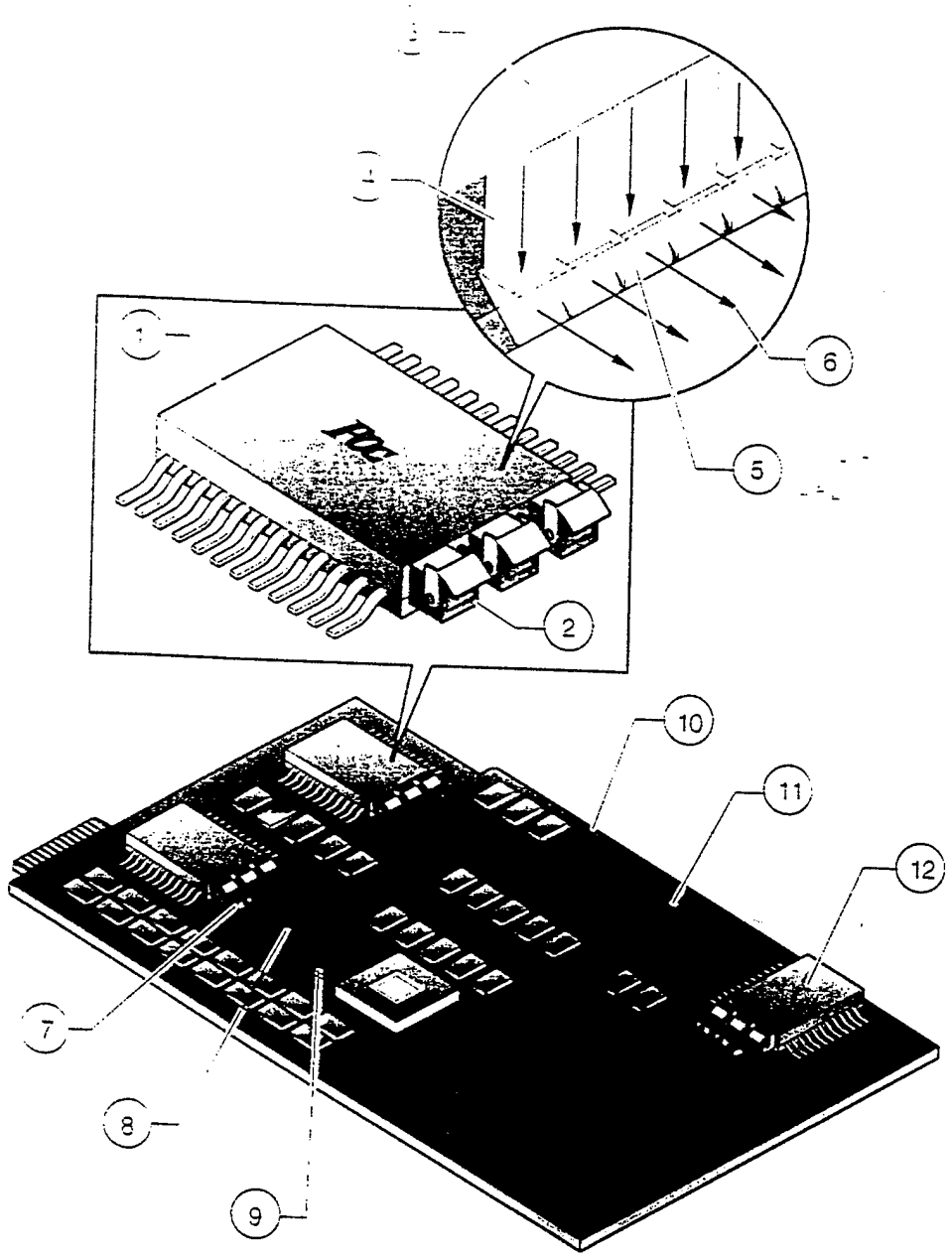


Figure 2-1
Multiboard system for high speed processor. Communication coupling is through high speed optical interconnects (data, address) and slower electronic interconnects (control, power supply).



Legend:

1., 12.	Chip with virtual optical I/O	6.	Coupled light
2.	Optical socket	7.	Optical connector
3.	Packaging	8.	Optical waveguides
4.	Coupling optics	9.	Electronic connections
5.	Prism couplers	10.	PC board
		11.	Electronic chips

Figure 2-2
Optical interconnect design. Laser diodes (LDs) and photodiodes (PDs) are mounted on top of the optical waveguides.

2.3 Electronic Encoder/Decoder Design

Recent advances in telecommunications have resulted in the development of new high speed electronic serial data transmission lines in the range of 1 Gb/s to 5 Gb/s. It has become feasible to multiplex a large number of electronic lines (~32-64 bit bus) using a wide bandwidth optical line. This solution becomes more economical than conventional parallel data links, especially when considering cost/performance, radio frequency interference (RFI) suppressers, bit-to-bit skew, savings in the use of board real estate, and larger fan-out potential.

The cost of the silicon to perform such multiplexing operations is dropping, and several manufacturers are showing new IC designs. New chip sets from Vitesse and TriQuint semiconductor can achieve very high data rates (over 1 Gb/s). The G-Taxi chip set from Vitesse can support a transparent high-speed serial link between two high performance parallel buses. The chip set performs parallel-to-serial and serial-to-parallel conversion, by means of an 8B/10B coding scheme and serial transmission rates up to 1.25 Gb/s. The chip set also has the capacity to multiplex up to 40 bus lines. Assuming a standard 32 bit data bus, the single bus speed can be as high as 33 MHz and still be carried on a single optical line (see Figure 2-3).

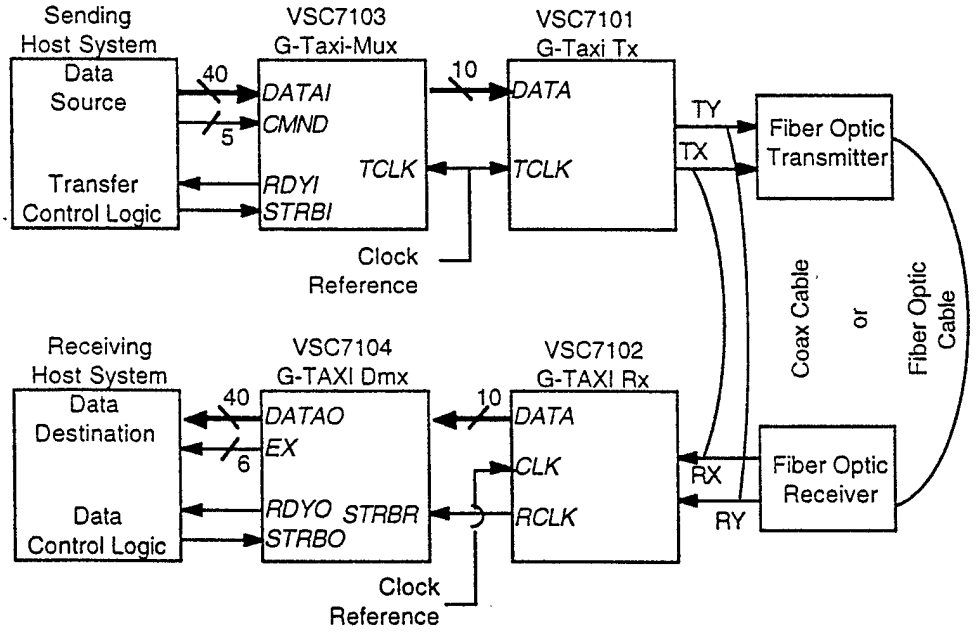
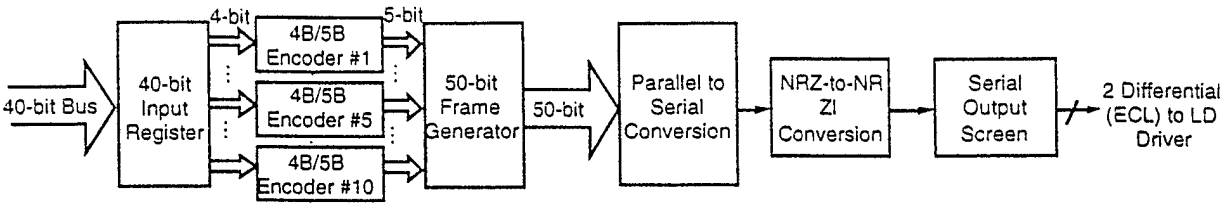
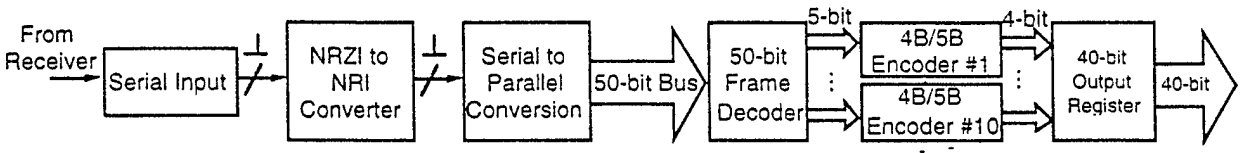


Figure 2-3
Optical communications, multiplexing on a 32 bit bus using G-Taxi chip set from Vitesse

Each G-Taxi chip set contains one transmitter and one receiver, and has a unidirectional 1.25 Gbit/s communication capability. The chip set can be considered as a 40 bit parallel register. The transmitter accepts a 40 bit word of TTL-level data, and then multiplexes and encodes it on a single serial output. The receiver decodes the data, performs serial-to-parallel conversion, and finally outputs the 40 bit parallel word on its TTL level bus. All of these operations, as well as the serial interface, are transparent to the system. The data transformations are shown in Figure 2-4(a) for the transmitter and 2-4(b) for the receiver.



(a) Transmitter.



(b) Receiver for serial optical communication

Figure 2-4
G-Taxi data transformations.

Other higher speed hardware includes the Vitesse VS804/805 and VS8021/VS8022 chipsets operating up to 2.5 Gbit/s.

2.4 Data Encoding and Decoding

The non-return to zero (NRZ) data stream must be encoded and decoded in order to perform clock recovery for high data rates. Most data transmission chip sets use codes such as 4B/5B, 8B/10B and 10B/12B. Here we present the basics of 8B/10B encoding, available on the Vitesse chip set.

8B/10B encoding is performed in the subblocks, 5B/6B and 3B/4B. A block code of 8 bits of user data is divided into 5 bit and 3 bit subblock words for serial transmission. The user data is divided into 4 bit (3 bit) subblocks, where each subblock is encoded into 5 bit (4 bit) NRZ code words. Each encoded subblock code word can have an equal number of ones and zeros (disparity = 0), more ones than zeros (disparity +2), more zeros than ones (disparity -2).

The running disparity rule requires that the 4B/5B codes encode the user data in such a way that

1. A code word with disparity = 0 is used to uniquely represent a data segment, and
2. A code word with disparity +2 and its binary complement (with disparity -2) are used together to represent one segment of data.

An 8B/10B encoding example is shown in Table 2-1:

Table 2-1. 8B/10B Encoding Example

Hex user data:	72	10
Binary user data:	01110010	00010000
Bit definitions	abcdefghij	abcdefghij
Initial RD	(-1)	
5B/6B Encode	010011 (-1)	
3B/4B Encode	1100 (-1)	
5B/6B Encode		01101 (+1)
3B/4B Encode		0100 (+1)
Final (RD)	(-1)	
Serial bit stream:	01001111000110110100	

The encoded serial data has 10 zeros and 10 ones, with recurring disparity (RD) = -1.

2.5 Experimental Setup and Results

In the experimental part of this project we have developed and tested an optical communication link that can be used for high speed data transfer. The transmitter module consists of three main units:

- Digital signal driver
- Laser diode driver
- Automatic power control unit.

The digital signal driver can be either a G-Taxi chip from Vitesse or a HotRod™ from TriQuint. Both support data rates in excess of 1 Gbit/s. The laser diode driver must perform high bandwidth current modulation. The LD driver, the schematic for which is shown in Figure 2-5, was then mounted on a printed circuit board (PCB). The transmitter module is shown in Figure 2-6. The transmitter prototype was mounted on a printed circuit board measuring 35 mm × 15 mm. Further size reduction by 50% is possible by using more compact components.

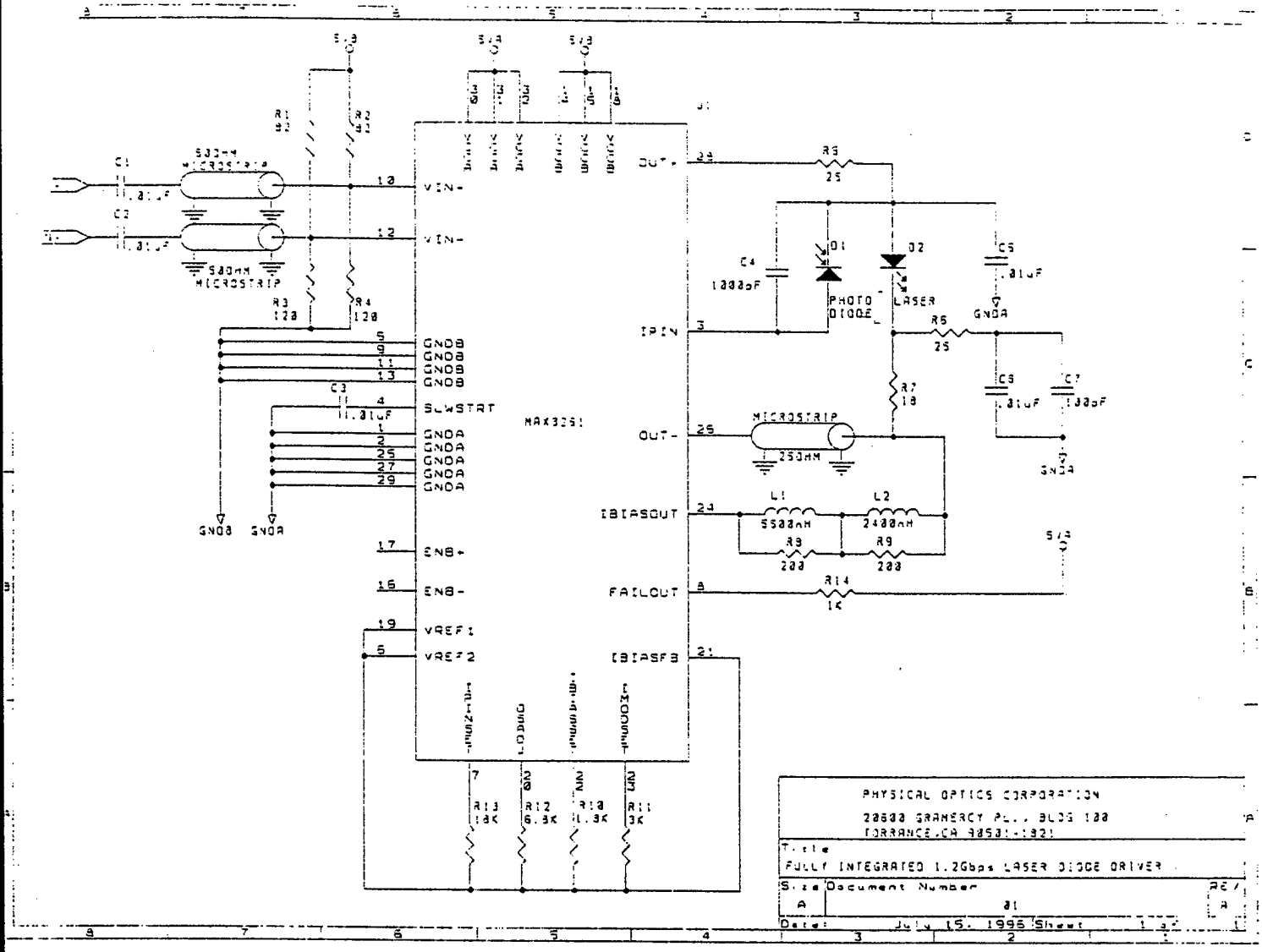


Figure 2-5
Schematic of the LD driver circuit with differential ECL signal compatibility.

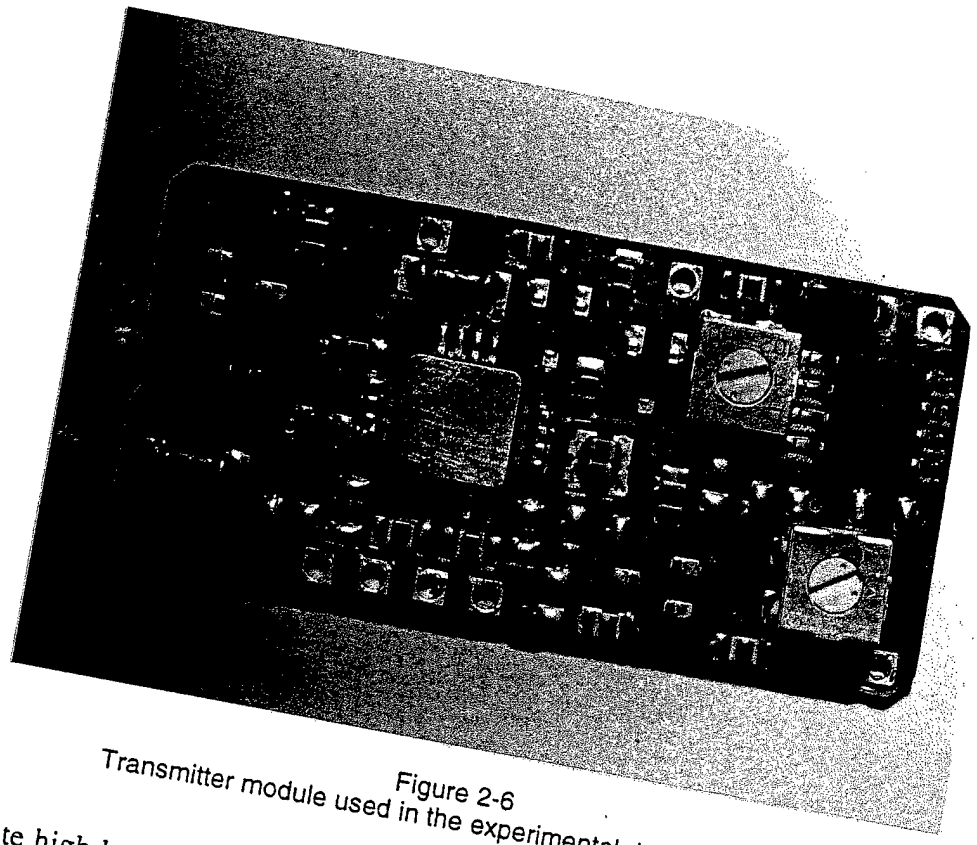


Figure 2-6
Transmitter module used in the experimental demonstration.

To demonstrate high bandwidth modulation capability, we used an HP-8133A pulse generator. The optical output was fed into a Tektronix SA-42 optical-to-electronic converter with over 2 GHz bandwidth. The converter output is connected to a high speed S-6 sampling head, which in turn is plugged into a 7904 high bandwidth Tektronix oscilloscope. In order to test the performance of the optical data channel, the transmitter and receiver were separated and a collimated LD beam was used to illuminate the detector.

The performance of the optical data channel was tested at data rates ranging from 600 Mb/s to 1.2 Gb/s. The oscilloscope trace for 600 Mb/s is shown in Figure 2-7 and for 1.2 Gb/s in Figure 2-8.

The relationship between the square wave modulation bandwidth and the data rate is that one period of the square wave signal corresponds to two data bits, (e.g., a 600 MHz square wave corresponds to a data rate of 1.2 Gb/s for NRZ binary signals.

Further testing included measuring changes in the output spectrum as a function of the modulation frequency. Figure 2-9 shows the wavelength distribution without any modulation, and modulated at 600 MHz and 1.2 GHz. Center wavelength shifts can be observed in those graphs.

Figure 2-7
Oscilloscope trace for 600 Mb/s optical modulation

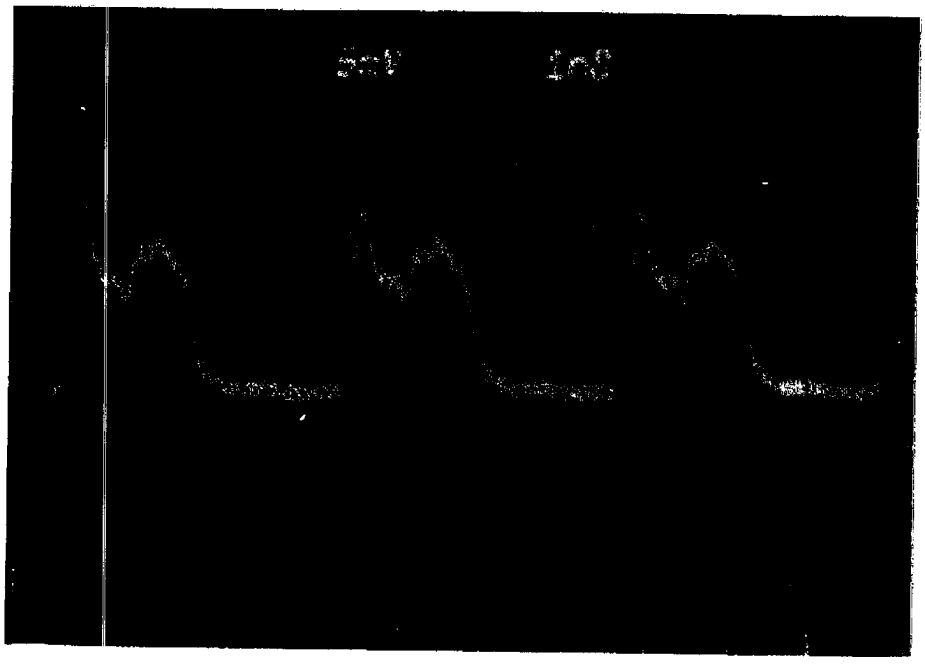
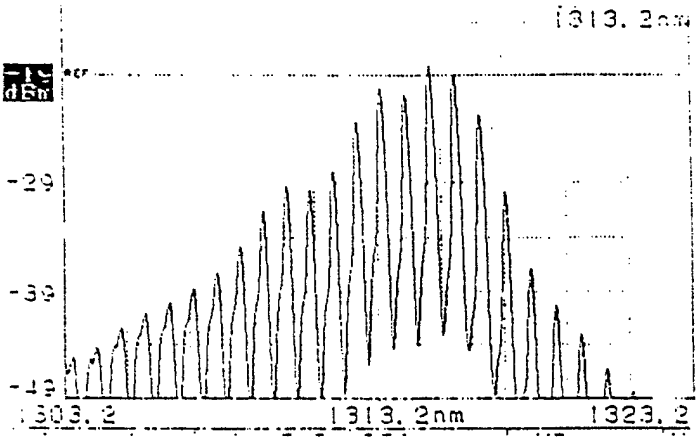
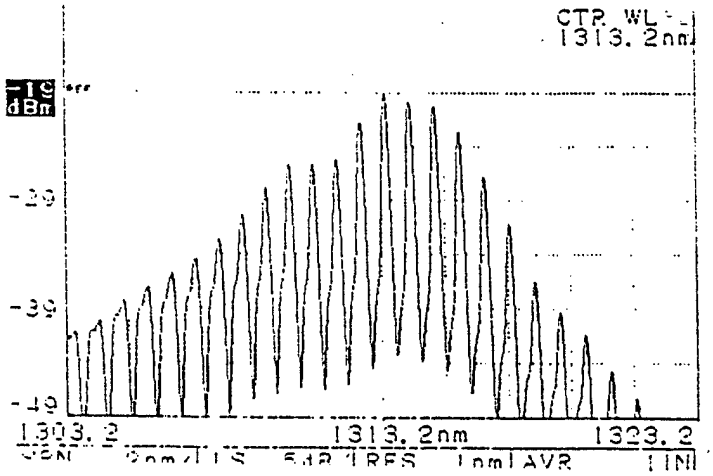


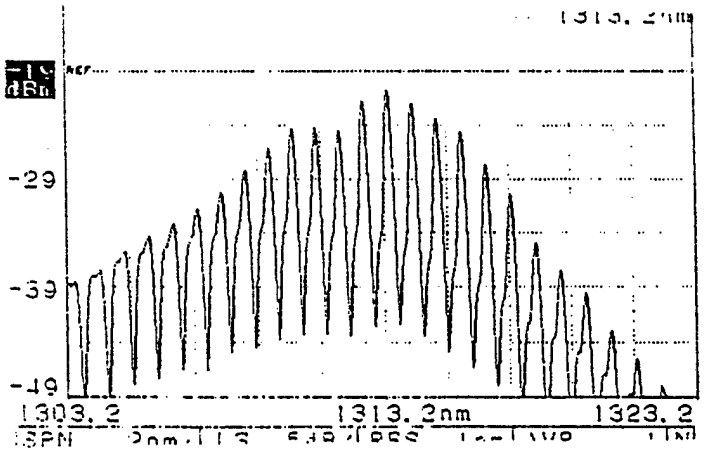
Figure 2-8
Oscilloscope trace for 1.2 Gb/s square wave signals transmitted over optical data channel.



(a) No modulation.



(b) 600 Mb/s modulation.



(c) 1.2 Gb/s modulation.

Figure 2-9
Optical output spectrum showing modulation shift towards shorter wavelengths with increased modulation speeds.

2.6 Bidirectional Optical Waveguide Interconnects

This section discusses key system design considerations such as optical parallel interconnection architecture, grating coupler efficiency analysis including fan-out/fan-in, waveguide array fabrication, optimization for the bidirectional optical data transfer, as well as limitations imposed by high speed electronics.

An optical multi-board communication system is sketched in Figure 2-10. A laser diode is bidirectionally coupled through a holographic grating coupler waveguide hologram and into the waveguide structure. The fan-out is implemented as a bidirectional holographic grating coupler that is illuminated by the laser. Each node contains both an LD and a photodiode associated with a transceiver that detects light from either hologram array. This is because the light travels in both directions in the waveguiding.

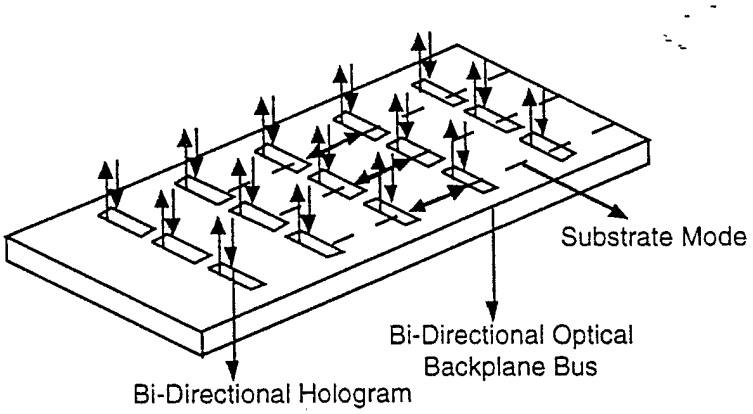


Figure 2-10
Bidirectional optical multi-board system.

The signal flow between the optical interconnect medium and the processor/memory boards must be bi-directional. Multiplexed waveguide holograms facilitate two-way communication between boards. In additional to the optical channels, the system includes lower speed electronic signals as well as power supply lines.

Optical signals enter the optical waveguides through a holographic grating coupler, which forms a totally internally reflected beam within the guiding plate. An array of fan-in/fan-out holograms are recorded at equal intervals corresponding to the positions of the electronic boards. The effect of the bi-directional waveguide array is to greatly accelerate the application that uses the transceiver system. Through the holographic bidirectional communication channels, each electronic board can send and receive information to and from every other board in the system through the fully interconnected optical communication network.

The sensitivity of a photodiode detector is a function of laser power, modulation speed, bit error rate, and the wavelength of the signal carrier. For a PIN-FET photodiode with a quantum efficiency of 50% at $\lambda = 1.3$ μm, a data transfer rate of 1.2 Gb/s, and a permissible error probability after amplification of the detected signal of less than 10^{-9}, the minimum modulated power required at the detector site is determined to be 5.08×10^{-5} W [1]. For the optimized bidirectional optical bus, simulation predicts an output efficiency of ~ 6.3%, which requires a minimum modulated input power of 0.8 mW.

2.6.1 Fabrication Considerations for the Optical Interconnections

The optical interconnections can be implemented as a thin waveguiding layer with an array of 1-D holographic grating couplers. The preferred material for implementing the waveguides is DuPont photopolymer deposited on the PCB substrate, because it permits the use of dry processing for fabrication of the holographic grating couplers for both fan-out and fan-in. The principle of operation of the grating couplers is based on the Bragg diffraction effect, which is characterized by high angular and wavelength dependence [2].

Light propagates through an optical waveguide in a series of diffraction (at input and output grating couplers) and reflection processes (propagation within the waveguide). The zig-zag guided substrate waves go through many iterations of this cascaded fan-out process until they hit the last holographic element.

The diffraction efficiency of the holographic grating couplers must be adjusted to produce uniform intensity at each surface-normal fan-out as a function of the diffraction efficiency of the corresponding holographic grating. By changing the diffraction efficiency distribution of the holographic grating arrays, we can manipulate the fan-out intensity distribution. It is impossible to achieve a uniform fan-out intensity distribution for all cases where the modulated optical signals are incident from different channels, because of the inherent bidirectionality of the optical interconnects. For example, a multi-processor system containing MCM modules requires n(n-1) interconnects to support the "broadcasting" function of the interconnection protocols. Each module must be interconnected with the other n-1 modules. A uniform fan-out intensity in the first module (which acts as the input module, and the rest as receiving ports) will make the power budget worse when this is reversed and the n^{th} module is treated as the input and all the others as the outputs. In other words, the optimal design should minimize power fluctuations rather than equalizing power distribution among n(n-1) interconnect scenarios.

A schematic of the bidirectional optical interconnects among five boards on one side of the substrate is shown in Figure 2-11. Here we assume that the diffraction efficiencies of the first hologram array are, from left to right, η_1, η_2, ... and η_N. Note that N = 5 in our case. To optimize the power budget, we must impose the following criterion: $\eta_1 = 1$ and $\eta_N = 0$; i.e., there is only one hologram at the first and the Nth channels. If we denote P_{ij} to be the output power at the jth channel when light is incident from the ith channel, and the same holds for P_{ji}, except that the direction of propagation is reversed (Figure 2-11), we have

$$P_{ij} = 0 \qquad (2\text{-}1)$$

whenever i = j,

$$P_{i1} = P_{iN} = 0 \qquad (2\text{-}2)$$

where i = 1, ..., N.

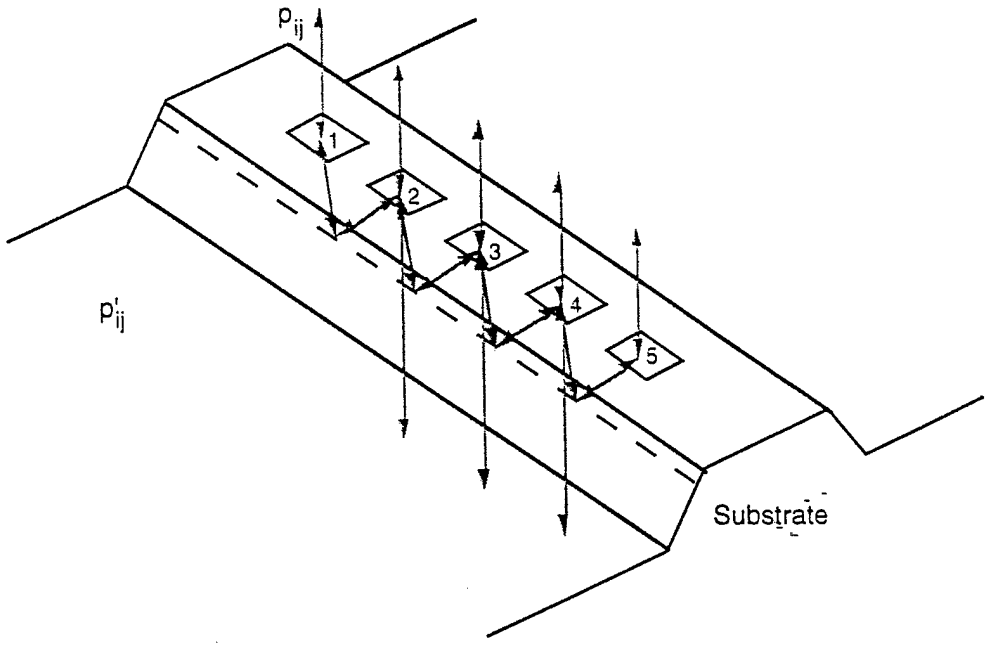


Figure 2-11
Five-channel bidirection optical interconnect concept.

Our optimization problem is to find a distribution of diffraction efficiencies that produces a fan-out intensity distribution with high uniformity, regardless of which channel the light is incident on. Optimizing the objective function subject to certain constraints, leads to a well balanced fan-out distribution. For our problem, an obvious objective function is the sum of the square value of the differences between the fan-out intensities and their average. The objective function is:

$$E = \sum_{i=1}^{M}\left[\sum_{\substack{j=1\\ j\neq 1}}^{N} W_{1ij}\left(\frac{P_{ij}}{P}-1\right)^2 + \sum_{j=2}^{N-1} W'_{1ij}\left(\frac{P'_{ij}}{P}-1\right)^2\right]$$

$$\sum_{i=1}^{M}\left[\sum_{\substack{j=1\\ j\neq 1}}^{N} W_{2ij}\left(\frac{\overline{P}}{P_{ij}}-1\right)^2 + \sum_{j=2}^{N-1} W'_{2ij}\left(\frac{\overline{P}}{P'_{ij}}-1\right)^2\right] \qquad (2\text{-}3)$$

where

$$W_{1ij}^{(')} = \exp\left[A\left(\frac{P_{ij}^{(')}}{\overline{P}}-1\right)\right] \qquad (2\text{-}4)$$

$$W_{2ij}^{(')} = \exp\left[B\left(\frac{\overline{P}}{P_{ij}^{(')}} - 1 \right) \right] \tag{2-5}$$

where $W_{1ij}^{(')}$ and $W_{2ij}^{(')}$ are weight factors, and $P_{ij}^{(')}$ is the general expression for the primed and unprimed power. A and B are iteration factors.

The square value of the difference between P_{ij} (or P'_{ij}) and $\overline{P}$ can be increased by multiplying each term of Eq. (2-3) by an appropriate statistical weight. This statistical weighting should give us a more nearly optimal result. After comparing all the results, we can find the optimal result. This idea can be grasped by assuming that $A \geq 0$ and $B \geq 0$, and alternately changing the values of A and B.

By minimizing the objective function E (see Eq. (2-3)) an optimized fan-out distribution can be found. In that optimal case the first derivatives of E with respect to η_i (i=2, ..., N-1) are equal to zero:

$$\frac{\partial E}{\partial \eta_2} = 0, \ \frac{\partial E}{\partial \eta_3} = 0, \ \frac{\partial E}{\partial \eta_n} = 0 \ . \tag{2-6}$$

We have optimized the fan-out distribution for the case in which N = 5. A computer program has been implemented that computes the fan-out intensities and their first derivatives. The three nonlinear equations (Eq. (2-6)) are then solved numerically using the Levenberg-Marquardt algorithm and a finite-difference approximation to the Jacobian algorithm [9-12].

The optimized diffraction efficiency distribution is calculated as

$$(\eta_1, \eta_2, \ldots, \eta_5) = (1.0, 0.3336, 0.2500, 0.4992, 0.0), \tag{2-7}$$

and the optimized fan-out intensities are given in Figure 2-12, in which P1 (j)s are our P_{ij}'s, and P2 (j)s are P'$_{ij}$'s. Because the average value of the fan-out intensities is 0.06667, we see that the minimum value of the fan-outs is very close to average. Furthermore, the maximum differs more the average, as a result of from the bidirectionality of the optical bus. Results of optimization are shown in Figure 2-12.

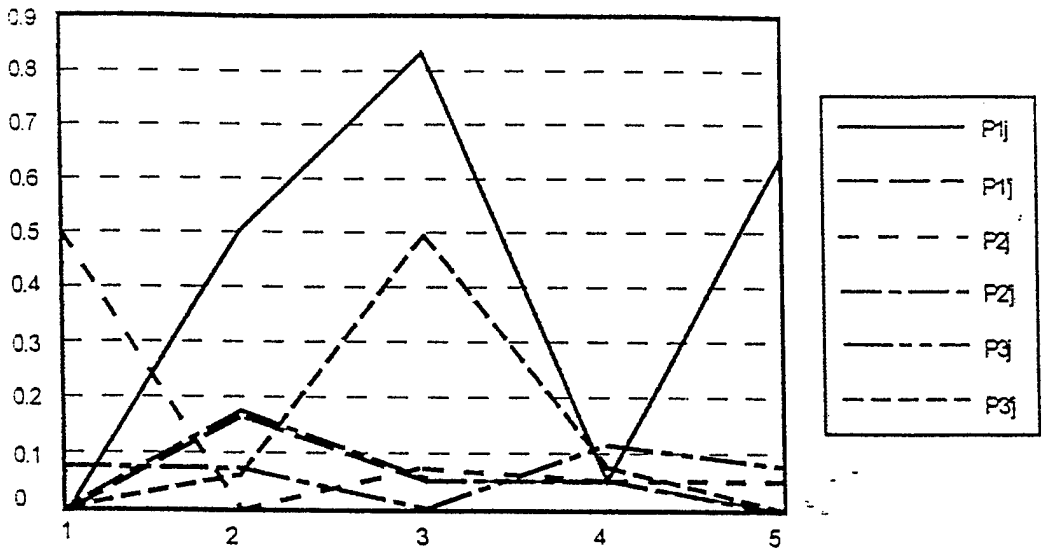


Figure 2-12
Distribution of optimized fan-out intensities.

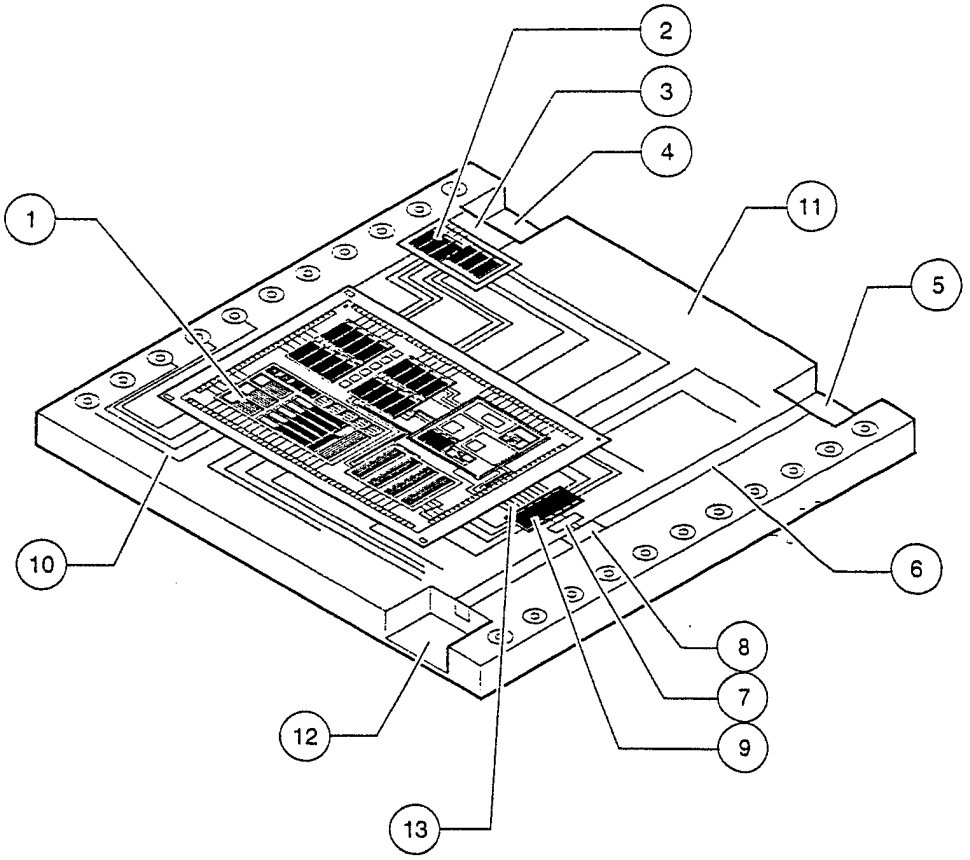
Optical interconnection systems with larger fan-outs and fan-ins can be optimized in a similar manner.

2.7 Design Issues

This combination of electronic processing with optical interconnections requires the integration of several components inside the new generation of ICs, as well as the development of special PCB integration techniques that include fiber optic modular tracks mounted directly on the PCB and special connectorization of ICs and optical fibers.

In order to show the feasibility of the concept, a prototype was designed that includes several discrete components in die form mounted on a single substrate (see Figure 2-13). These can be mounted using either flip-chip or wire bonding techniques. These components were: an electronic processor die with a large number of I/O pins for data and address bus, an optical virtual pin driver functioning as a parallel-to-serial shift register, and an EO modulator driver mounted on a channel waveguide.

The IC has both standard electronic metal (physical) pins and two or more optical virtual pins that are equivalent to a large number of physical electronic pins. This can be accomplished by multiplexing a large number of electrical signals from an entire 128-bit wide data bus onto a single optical fiber. *This technology directly reduces the number of I/O pins, which is one of the major bottlenecks in present and projected ICs.* The pin compression ratio can be as high as 100:1, or even 200:1, depending on the bandwidths of the multiplexed signals. The signals from the processing chip -- (1) in Figure 2-13 -- are multiplexed using a shift register (9) to produce serial data streams. Assuming a 128 data-bit bus operating at 10 Mb/s, the entire bus can be multiplexed onto a single optical fiber carrying a 1.28 Gb/s serial data stream.



1. Processing chip socket
2. Serial-to-parallel converter (demultiplexer)
3. PIN photodiode
4. Receiving optical port (virtual receiving pin)
5. Output optical port (virtual transmitting pin)
6. Channel waveguide
7. EO modulator driver
8. EO modulators with electrodes
9. Parallel-to-serial converter
10. Connections to standard electronic pins
11. Carrier substrate
12. CW laser input port

Figure 2-13
Top view of the virtual optical pin chip, including electronic and optical components.

The major component of the virtual optical pin (VOP) is the external EO modulator (8), which converts CW laser light entering through the input optical port (12) into a corresponding optical signal modulated at high speed. This modulator can be integrated with either a single channel waveguide (as a part of the waveguide structure) or a Mach-Zehnder interferometric modulator containing two arms with electrodes. In both cases, the EO modulators are integrated with the waveguides by means of standard VLSI technology. The EO modulator electrodes are wire bonded to the modulator driver. The third optical port (4) receives the data from off chip. The optical signals are converted to electrical signals through the PIN photodiode (3). The signals from the photodiode are conditioned (amplified, filtered, and clock recovered) before final demultiplexing through a serial-to-parallel shift register.

The conventional IC pins (10) for ground, power, and control signals will be exactly the same as in standard IC packages.

Another design issue is the mounting of the VOP-based IC on a PC board. Instrumenting the VOP will require special connectorization between the IC and optical fiber. The overall packaging of the

VOP chip also includes the fiber optic link between ICs through flexible fiber optic modules connectorized for easy installation. Four fiber optic modules are used:

- Fiber optic-IC connectorization module
- Straight wire module
- 45° module
- Y-junction module.

These modules are shown in Figure 2-14. All of these modules offer designers full flexibility for high interconnectivity among VOP ICs with minimum real estate requirements.

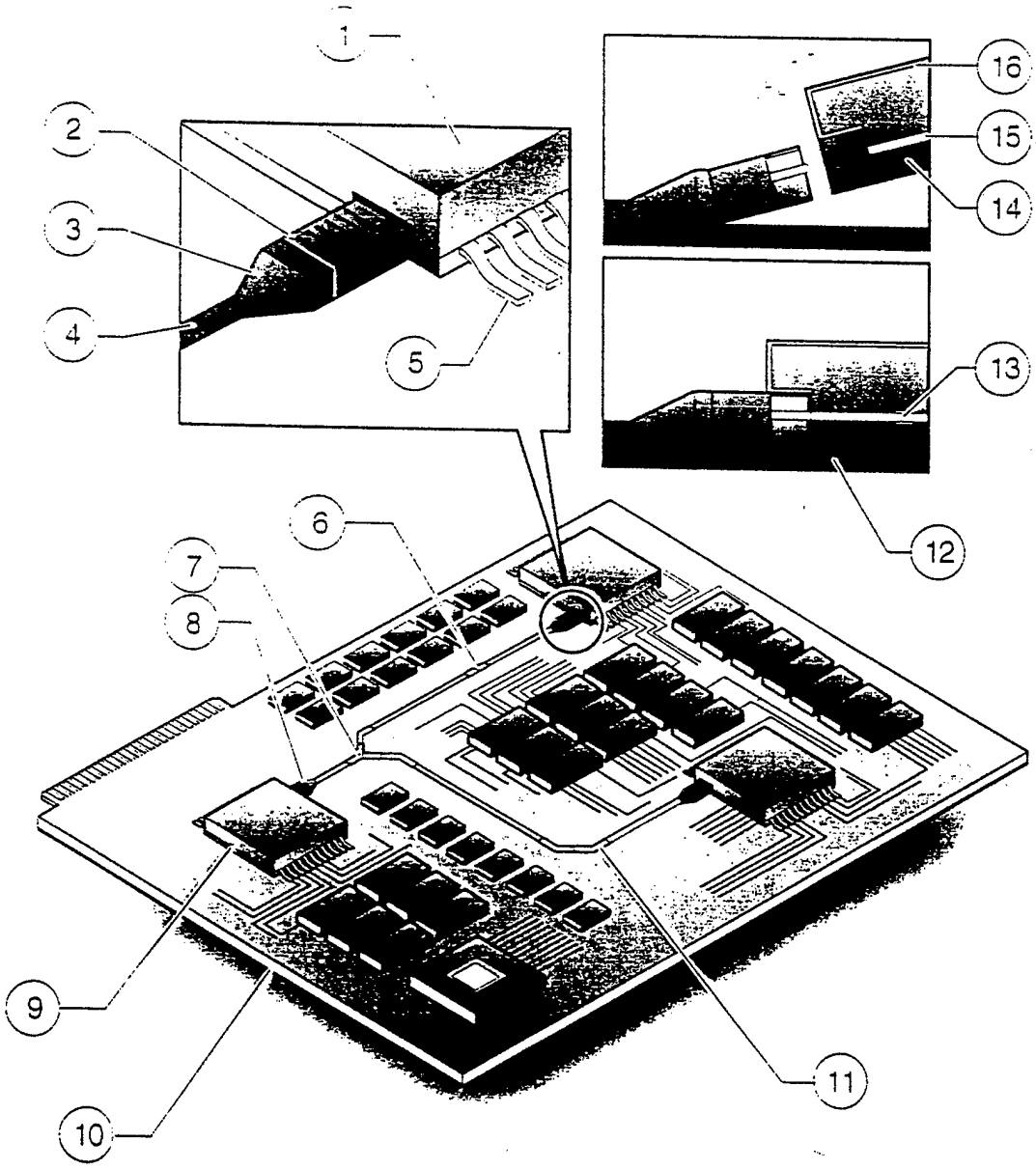


Figure 2-14
Modular approach to fiber optic connectorization.

2.8 Computer Simulation of Channel Cross-Coupling in a Single-Mode Optical Waveguide Bus Array

Wave propagation and cross talk in the waveguide bus array were simulated in two dimensions using BMP_CAD software. A five-channel bus array with the refractive distribution shown in Figure 2-15 was used with following waveguide parameters:

- Initial waveguide width — 2 μm
- Initial separation — 8 μm
- Complex refractive index — (1.5, 0.0)
- Wavelength — 0.6328 μm
- Wafer length — 900 μm
- Wafer width — 50 μm
- Base refractive index — (1.49, 0.0)
- Gaussian input field
  - centered at — 8 μm, -8 μm
  - halfwidth — 1.0 μm

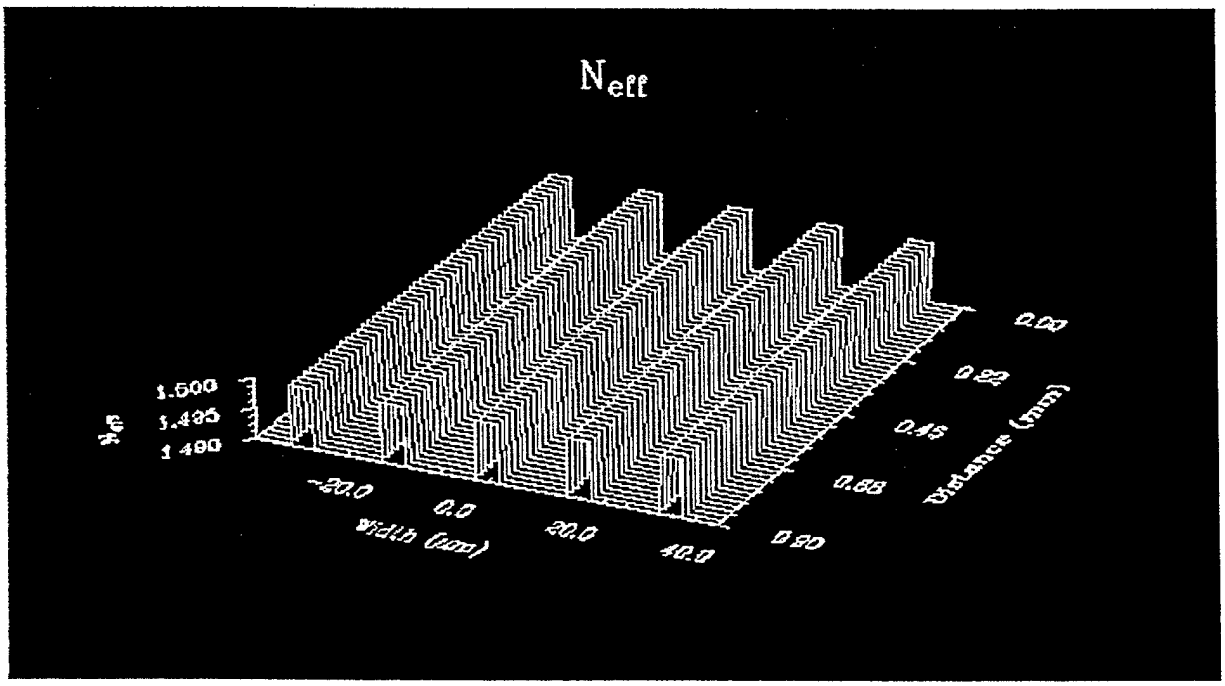


Figure 2-15
Distribution of the refractive index of the bus array.

The initial and boundary conditions were chosen to correspond to a symmetric Gaussian optical field ($\lambda = 0.6328$ μm) coupled into the two channels adjacent to the central waveguide, as shown in Figure 2-16. A paraxial Fresnel approximation of the Helmholtz equation was used in the calculations of the wave propagation through the waveguide array. Figure 2-17 shows the field distribution at the device output. The calculation shows that cross talk does not exceed –33.9 dB: the ratio of power in the excited channels to the power in the central channel (noise) is 2447. Figure 2-18 is a topographical view of the field in the bus array and a 3-D plot of the field distribution

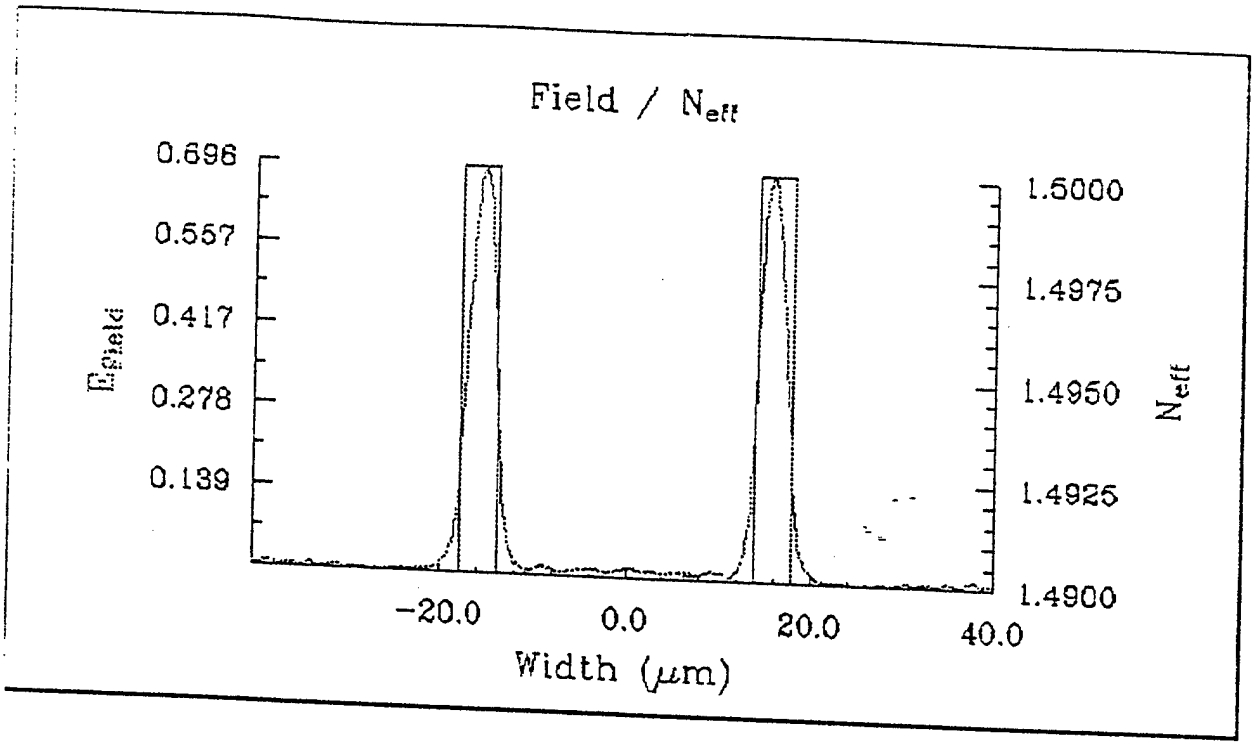


Figure 2-16
Input field distribution.

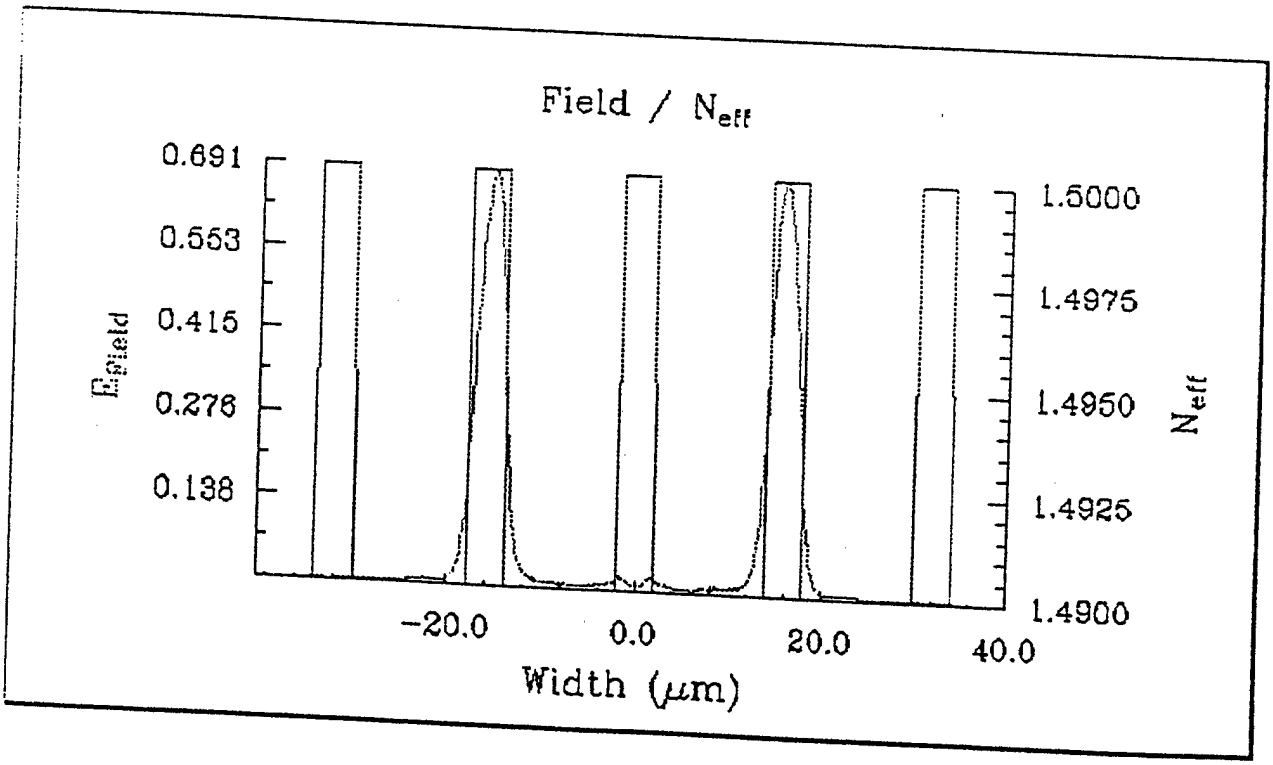
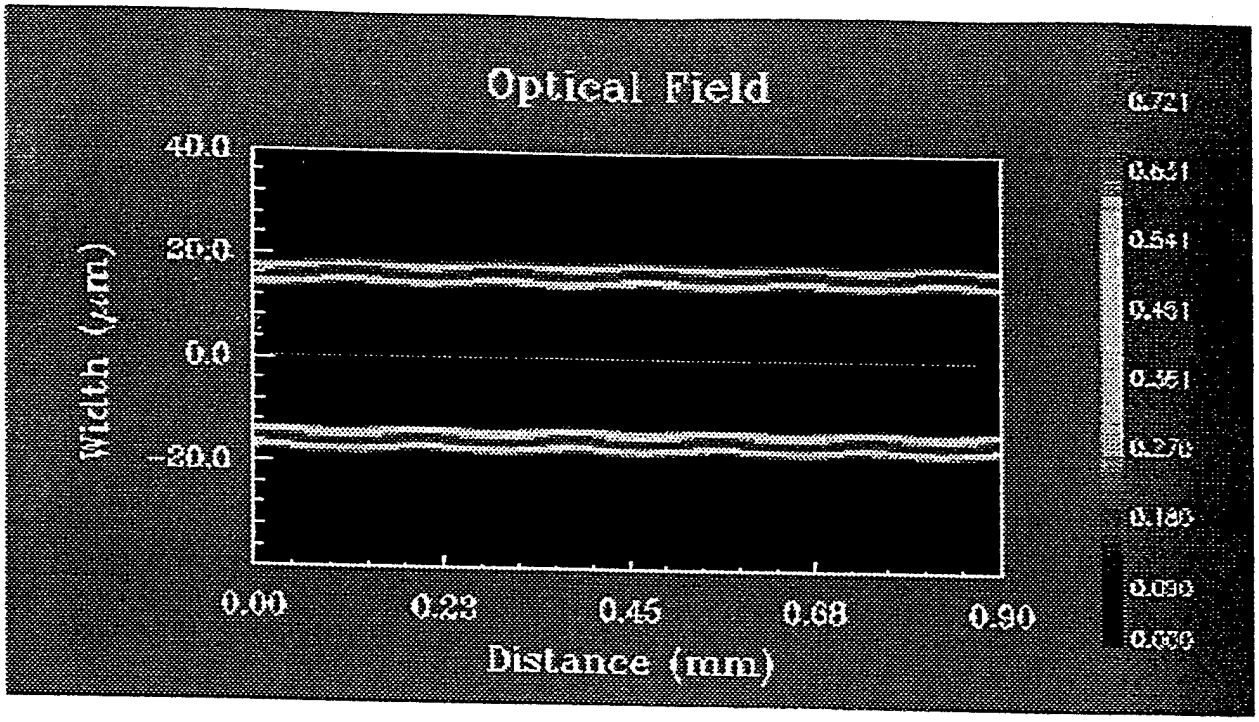
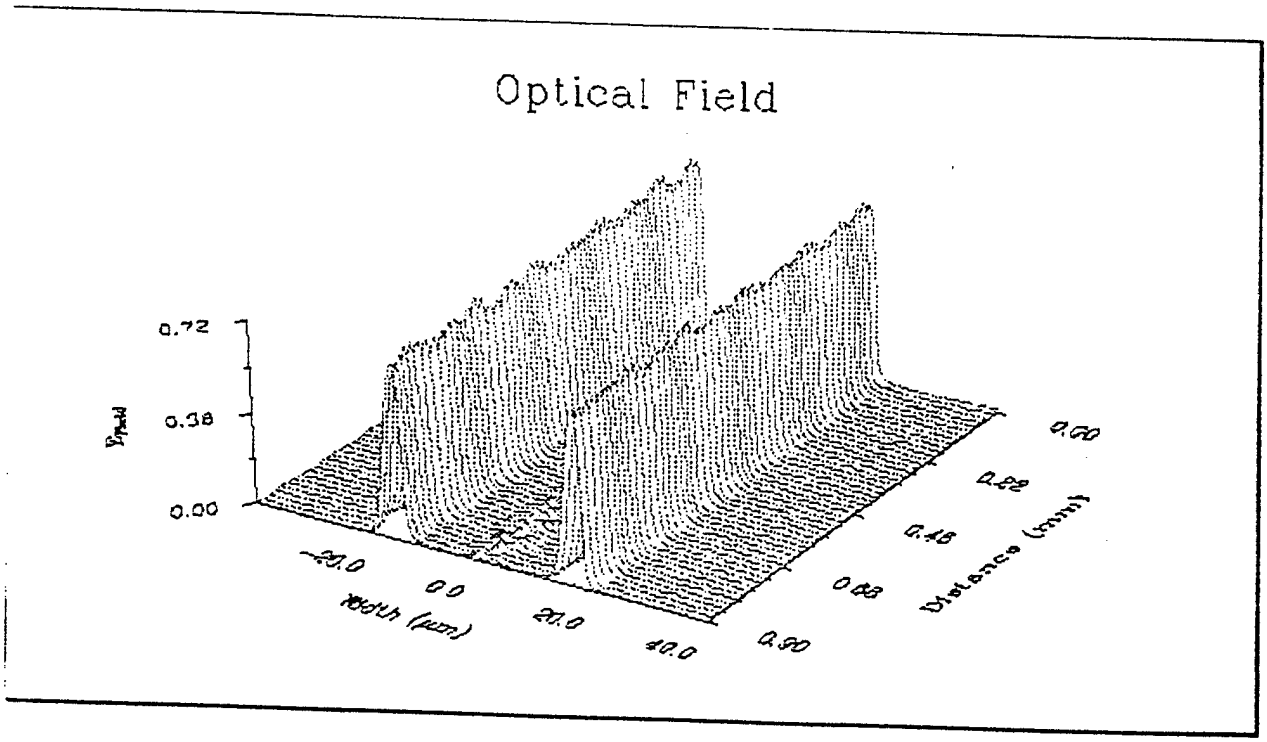


Figure 2-17
Output field distribution.



(a) Topographical view.



(b) 3-D graph.

Figure 2-18
Field distribution in the waveguide bus array.

2.9 Design of Multiprocessor System

The initial graphics accelerator board designed in this project had a single node architecture (see Figure 2-19). The functions of each building block of the board are discussed in turn.

In the multiple-node, four processing nodes are used in an MIMD architecture. In the single node prototype design, we use FPGAs to implement the graphics functions and other control units. VRAM and SRAM are used as main memory, instruction memory, look-up table and processor/display interface. The nodes are linked by means of optical connections.

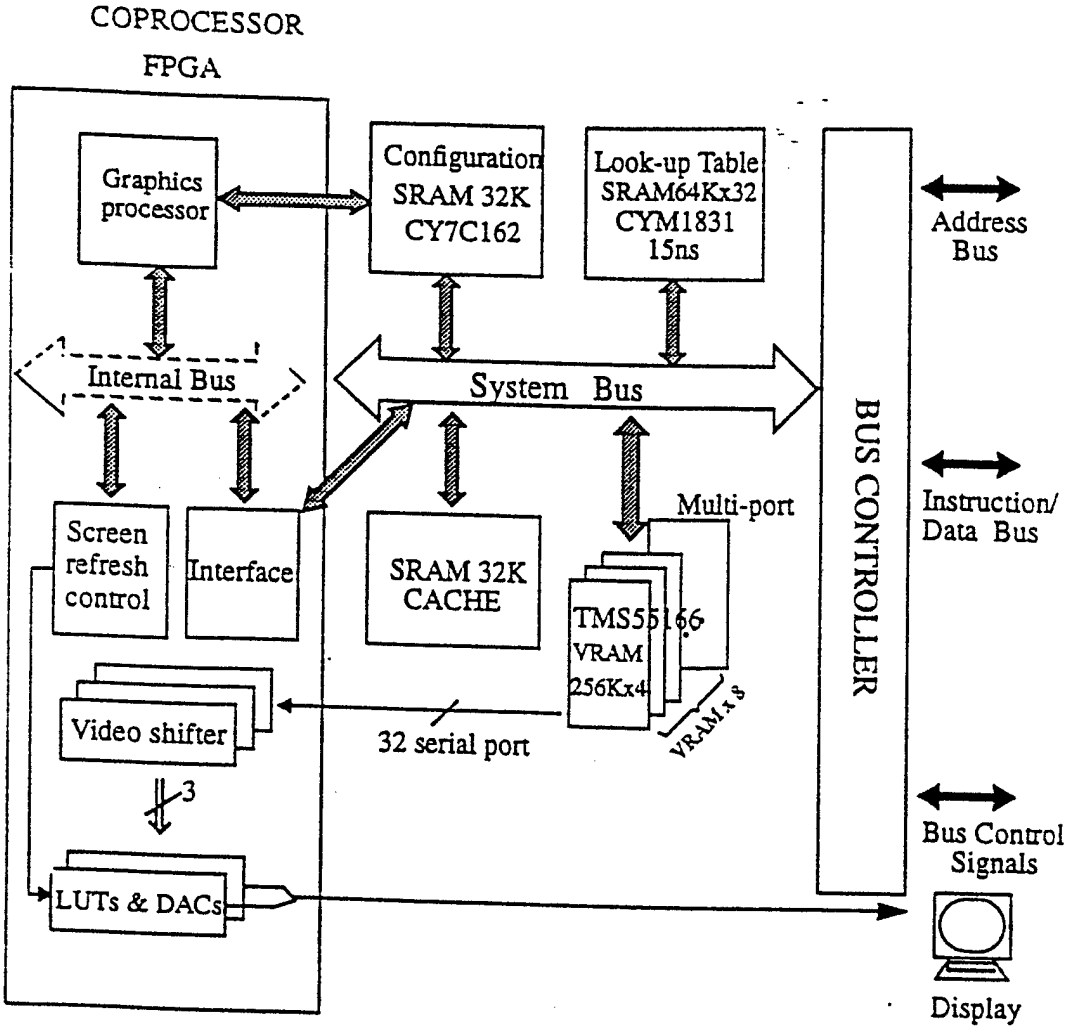


Figure 2-19
Single processing node architecture. Function-specific operations are stored in configuration SRAM. Fast look-up table operations are stored in 64k x 32 SRAM. VRAM can be accessed by the processor or by the look-up table processor. An FPGA will implement key functions of the TMS 34020 in a more efficient design.

The basic structure of this architecture is as follows:

- CPU: RISC embedded processor with on-chip floating-point unit (FPU)
- Coprocessor: Three reconfigurable FPGAs with their respective configuration SRAM or ROM
- System buses: Address bus, data bus, and control lines
- Bus extensions: Fast Serial Link (FSL) interface to host computer and Photonic Inter-Board Communication (PIBC) interface to other graphics processor boards
- Memory module: SRAM cache, DRAM, and memory controller

The RISC CPU performs the main computation tasks of the graphics algorithm. First the host computer generates executable code from the graphics algorithm and partitions the task among the graphics processor boards. Then the host computer transmits graphics data and programs to the processor boards and stores the data and programs into the memory module of each board. The RISC CPU can be idle when the host computer is accessing the memory module or reconfiguring the FPGA coprocessors. After this setup phase, the RISC CPU coordinates the three FPGA coprocessors through address decoding and interrupts. It can also access the memory module on another processor board through the Photonic Inter-Board Communication (PIBC) interface.

The three reconfigurable FPGAs work as coprocessors of the RISC CPU. They can be reconfigured as three separate coprocessors or as one corporate coprocessor. The configuration data is generated by the host computer and sent to the processor board through a Fast Serial Link (FSL) interface. This data is then stored in the accompanying configuration SRAM and is used to configure the FPGA chips for specific operations. The main function of these three FPGA coprocessors is floating-point geometric computation. They are accessed by the CPU through address decoding. The instructions and operands for the coprocessors are sent by the CPU through the data bus. After an operation, an FPGA coprocessor interrupts the CPU and sends the results to it.

The system buses, i.e., the address bus, data bus, and control lines, carry the signals between the RISC CPU, the FPGA coprocessors, and the memory modules. They also extend to other graphics processor boards and the host computer through the PIBC and FSL interfaces.

The FSL interface transfers the synchronization signal from the host computer so that every processor board in the network works in synch. Also, the FSL interface carries graphics data, programs, and configuration data between FPGAs and the host computer, and carries handshake signals to the host computer. The PIBC interface connects all the graphics processor boards into a ring network. Each board in the network has full access to other boards. The PIBC interface also gives the RISC CPU access to the memory module of the other boards, and transfers control signals between the CPU and another board.

The memory module consists of SRAM cache, DRAM, and a memory controller. SRAM is fast memory, used as the instruction and data cache for the RISC CPU and the three FPGA coprocessors. DRAM is the main memory on the processor board. It stores the graphics data and processing programs sent from the host computer. The refresh cycle of the DRAM is generated by the memory controller, which also allocates SRAM cache to the CPU and the three FPGAs. It also controls the timing of data flow between cache memory and DRAM to ensure data correctness.

2.9.1 Field Programmable Gate Array

In this prototype design, we implement the key functions of the TMS34020 and TMS34082 in an FPGA. The Texas Instruments TMS34020 is a single-chip graphics processor designed to accelerate 2-D displays on PCs and workstations. It can be paired with a TMS34082 floating-point

coprocessor, which speeds up the 3-D geometric transformations and clipping needed for 3-D graphics.

The FPGA implementation has several advantages:

- The main purpose of this design is to accelerate performance of the 3-D algorithms. Using an FPGA can achieve more customized functions.
- Many algorithms now run in software can be implemented directly in hardware. Because this kind of hardware is designed according to the needs of the algorithms, the delay is much less and the throughput rate much higher compared with the graphics processor-based design.
- Using an FPGA makes it much easier to construct a pipeline and/or parallel architecture.

The major computation-intensive operations realized by the FPGA are the floating-point operations for 3-D graphics and the common 2-D graphics operations. The 3-D operations are: calculating multiple linear interpolations in parallel, calculating multiple z-buffer comparisons in parallel, and conditionally updating multiple pixels in parallel. The common 2-D operations include copying pixels of 1 to 8 bits, automatically aligning, masking, and clipping, filling a rectangular array of pixels with a solid color, drawing a 1-pixel-thick straight line using a midpoint scan-conversion algorithm, and drawing the pixel indicated by a pair of (x,y) coordinates.

The other building blocks implemented by the FPGA are control units, video shifters, screen refresh control unit, DACs, and on-chip registers. The video shifters and screen refresh control unit are used when the results of a node can be sent to a display unit directly.

2.9.2 Video Random Access Memory

We use video random access memory (VRAM) as the main memory of this operating node, and SRAM both as the cache between the main memory and processor and as the look-up table.

A VRAM chip, as shown in Figure 2-20, is similar to a conventional DRAM chip but contains a parallel-in/serial-out data register connected to a second data port. Here the parallel port is used for data transfer between nodes and between main memory and cache memory. The serial port is useful when the results can be displayed directly.

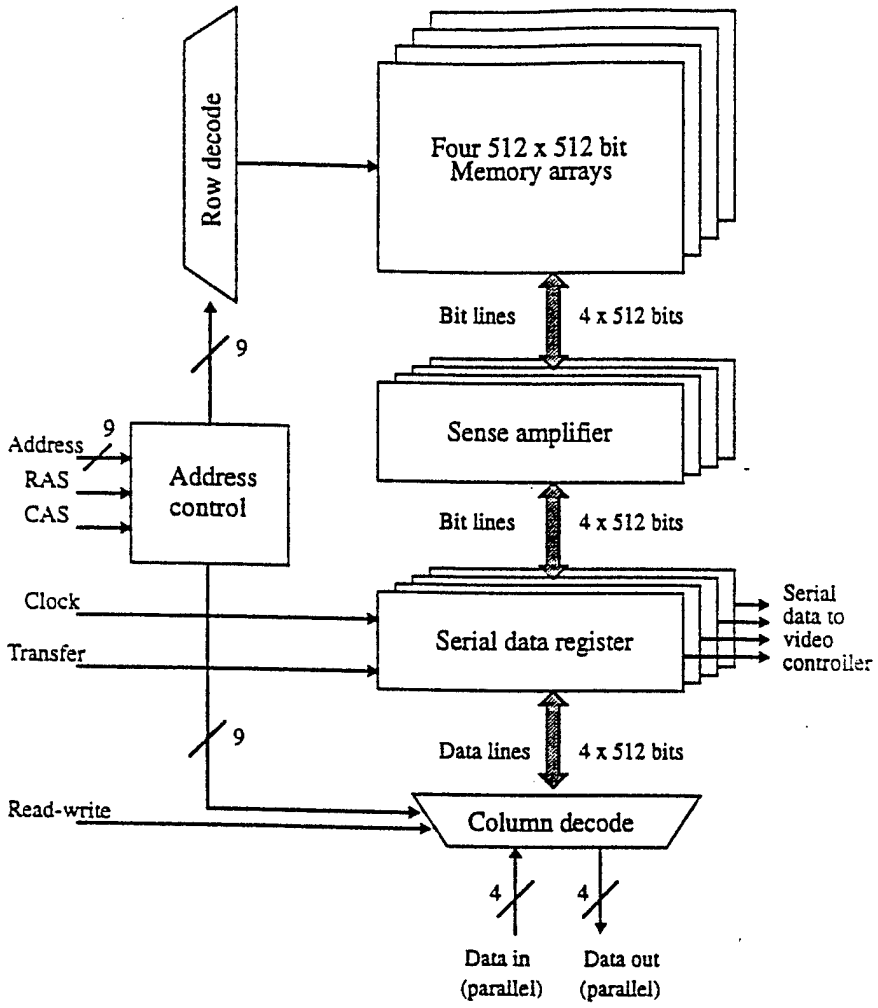


Figure 2-20
Diagram of a 1 Mbit (256K x 4) VRAM chip. The serial data register provides a second (serial) port to the memory array.

2.9.3 Static Random Access Memory

One 32k SRAM is used for instruction memory and another for cache memory, and a 64k x 32 SRAM stores the contents of the look-up table used by the graphics algorithms.

2.9.4 Implementation of Single Node Architecture in Parallel and Pipeline Structures

This section describes the implementation of a single node architecture in parallel and pipeline structures. Simulation result for each architecture are compared with those of existing graphics processors.

A number of single node processing architectures are appropriate for either a parallel SIMD architecture or a pipeline MIMD architecture to accelerate computation-intensive graphics algorithms.

2.9.4.1 Parallel Architecture

For a given image size, we first determine how many calculations must be performed serially and how many can be performed in parallel. By evenly distributing the parallel computations to each operating node and assuming that a given operation takes the same amount of time when running on an existing graphics processor as when running on the architecture we present here, we can determine the speedup achieved by this architecture.

Assume that α percent of the calculations must be done serially and the number of nodes used is n; then the speedup is

$$\text{Speedup} = \frac{T_{old}}{T_{new}} = \frac{1}{a + \frac{(1-\alpha)}{n}} \quad . \tag{2-8}$$

Using Eq. (2-8), we can find an optimal value for n.

2.9.4.2 Pipeline Architecture

The commonly used pipeline architectures for 3-D graphics are local illumination pipelines and global illumination pipelines. Several pipeline structures are shown in Figures 2-21 through 2-25.

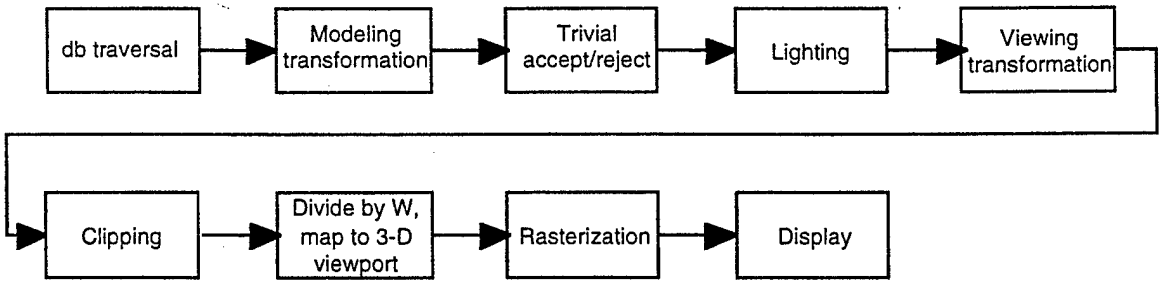


Figure 2-21
Rendering pipeline for Z-buffer and Gouraud shading.

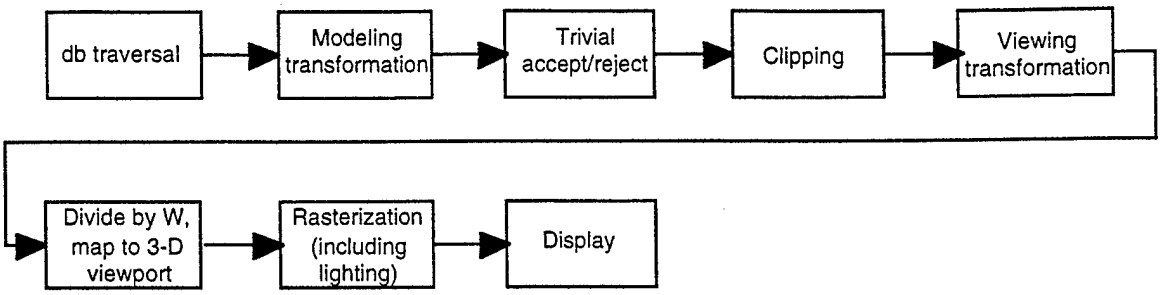


Figure 2-22
Rendering pipeline for Z-buffer and Phong shading.

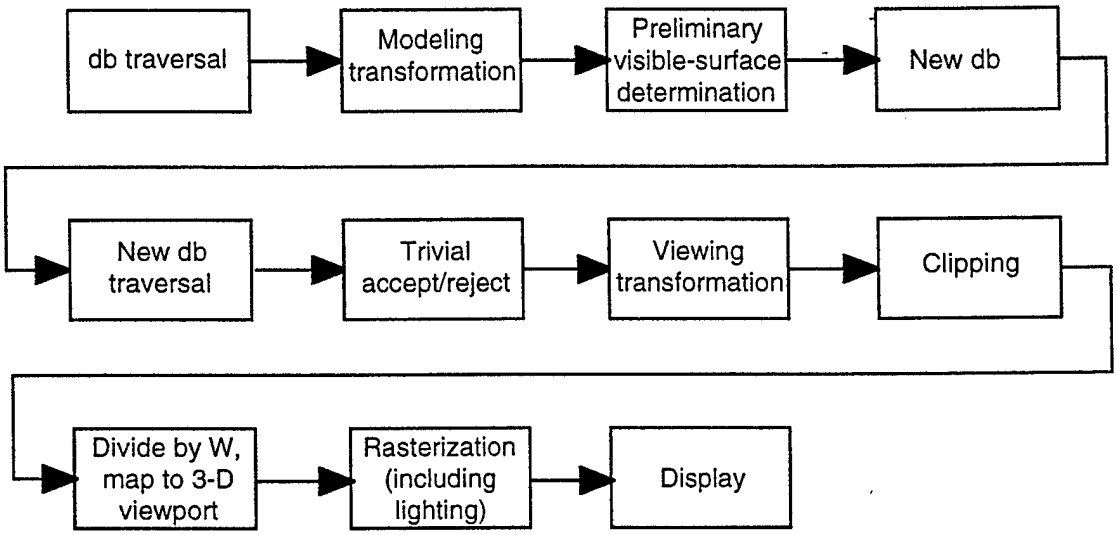


Figure 2-23
Rendering pipeline for list-priority and Phong shading.

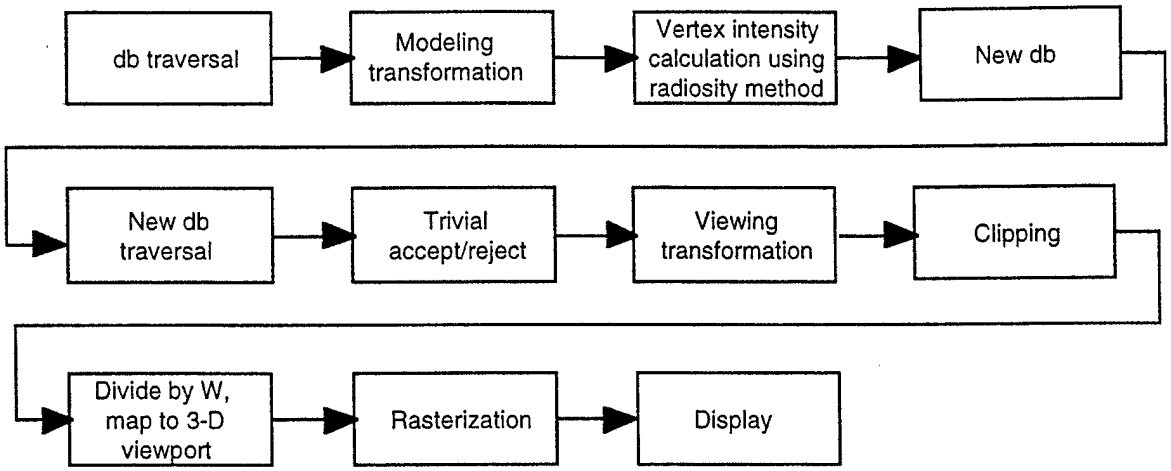


Figure 2-24
Rendering pipeline for radiosity and Gouraud shading.

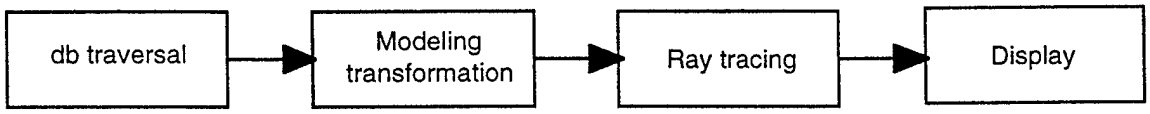


Figure 2-25
Rendering pipeline for ray tracing.

In our simulation, the pipeline is repartitioned at the very beginning according to the number of nodes given to make sure that the computation times of the stages are as similar as possible. Similarly, we assume the same operating speeds for the same duration when running on this kind of structure as when running on an existing graphics processor. Then, for a given image size we can analyze the delay time and the throughput rate of the pipeline structure. By changing the number of nodes we get the curves of delay time and throughput rate vs. number of nodes. From this we can determine the optimal number of nodes to use.

Let us take an example with which to evaluate the architecture. We assume that an ambient/diffuse illumination model and Gouraud shading are to be applied to each primitive. We assume a screen size of 1280 by 1024 pixels and an update rate of less than 10 frames per second.

In summary, our sample application has the following characteristics.

- 10,000 triangles (none clipped)
- Each triangle covering an average of 10 pixels, half being obscured by other triangles
- Ambient and diffuse illumination models
- Gouraud shading
- 1280 by 1024 display screen, updated at less than 10 frames per second.

We cannot precisely calculate all the computation and memory-bandwidth requirements in our sample application, since many steps are difficult to categorize. Instead, we concentrate on the three performance barriers: the number of floating-point operations for geometry computations,

the number of integer operations for computing pixel values, and the number of frame-buffer accesses for rasterization.

For every algorithm we apply for simulation, we can also compare these two kinds of architectures to see which one is more suitable for which algorithm. From this, we can find a better solution combining the advantages of the two.

Here, we can use a geometry calculation as an example. First, we assume that the execution time for multiplications and additions is the same, and that each single processing node can process 3 multiplications and 3 additions at one time. For each frame, we must process $10,000 \times 3$ vertices and vertex-normal vectors. In the modeling transformation stage, transforming a vertex (including transforming the normal vector) requires 25 multiplications and 18 additions. The requirements for this stage are thus $30,000 \times 25 = 750,00$ multiplications and $30,000 \times 18 = 540,000$ additions. Thus, the total execution time for the modeling transformation stage is $(25 \times 30,000)/(3 \times N)$, where N is the number of processing nodes.

Trivial accept/reject classification requires testing each vertex of each primitive against the six bounding planes of the viewing volume, a total of 24 multiplications and 18 additions per vertex. The requirements for this stage are thus $30,000 \times 24 = 720,000$ multiplications and $30,000 \times 18 = 540,000$ additions. The total execution time for this stage is $(24 \times 30,000)/(3 \times N)$ processor cycles.

Lighting requires 12 multiplications and 5 additions per vertex, so that the total execution time would be $(12 \times 30,000)/(3 \times N)$. The viewing transformation requires 8 multiplications and 6 additions per vertex, so that the execution time for this stage is $(8 \times 30,000)/(3 \times N)$.

The requirements for clipping vary; the exact number depends on the number of primitives that cannot be trivially accepted or rejected, which in turn depends on the scene and the viewing angle. We have assumed the simplest case for our database, that all primitives lie completely within the viewing volume. If a large fraction of the primitives must be clipped, the computation required could be substantial (perhaps even more than in the geometric transformation stage.)

Division by W requires 3 divisions per vertex, a total of $30,000 \times 3 = 90,000$ divisions. Mapping to the 3-D viewport requires 2 multiplications and 2 additions per vertex, for a total execution time of $(2 \times 30,000)/(3 \times N)$.

For rasterization calculations and frame-buffer accesses, we assume that z values and RGB triples each occupy one word (32 bits) of frame-buffer memory. For each pixel that is initially visible, values for z, R, G, and B are calculated (r additions per pixel if forward differences are used), a z value is read from the frame buffer (1 frame-buffer cycle), the z values are compared (1 subtraction), and new z values and colors are written (2 frame-buffer cycles). For each pixel that is initially not visible, only the z value need be calculated, and a z value is read from the frame buffer (1 frame-buffer cycle), and the two z values are compared (1 subtraction). Assuming that half of the pixels of each triangle are visible in the final scene, a reasonable guess is that three-quarters of the pixels are initially visible and the other quarter are invisible. Each triangle covers 100 pixels. Thus, to display an entire frame requires a total of $(750,000 \times 5)+(250,000 \times 2)=4.25$ million additions and $(750,000 \times 30)+(250,000 \times 1)=2.5$ million frame-buffer access is required. To initialize each frame, both color and z-buffers must be cleared, requiring an additional $1280 \times 1024 \times 2 = 2.6$ million frame-buffer accesses. The total number of frame-buffer accesses per frame, therefore, is 2.5 million + 2.6 million = 5.1 million.

Table 2-2 summarizes the floating-point requirements for all of the geometry stages and the rasterization calculations and frame-buffer accesses.

Table 2-2 Summary of the Calculation Needed for 3-D Display

Feature	Multiplication/Division	Addition/Subtraction	Total Execution Time (Processor Cycles)
Modeling Transformation	750,000	540,000	750,000/(3×N)
Trivial Accept/Reject	720,000	540,000	720,000/(3×N)
Lighting	360,000	150,000	360,000/(3×N))
Viewing Transformation	240,000	180,000	240,000/(3×N)
Divide by W	90,000	—	60,000/(3×N)
Mapping to 3-D View	60,000	60,000	60,000/(3×N)
Rasterization	—	4,250,000	4,250,000/(3×N)

Assuming the number of nodes is n, the performance is shown in Figure 2-26.

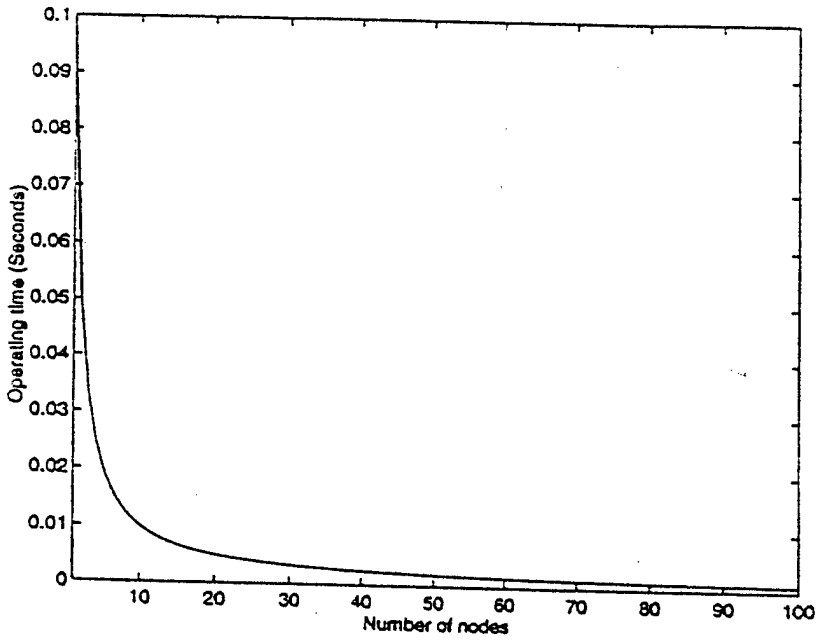


Figure 2-26
Operating time using a pure parallel architecture.

Considering the other operations related, we assume that 20% of operations must be done serially, so the performance is as shown in Figure 2-27. Figure 2-28 is the result, including both parallel and pipeline architectures.

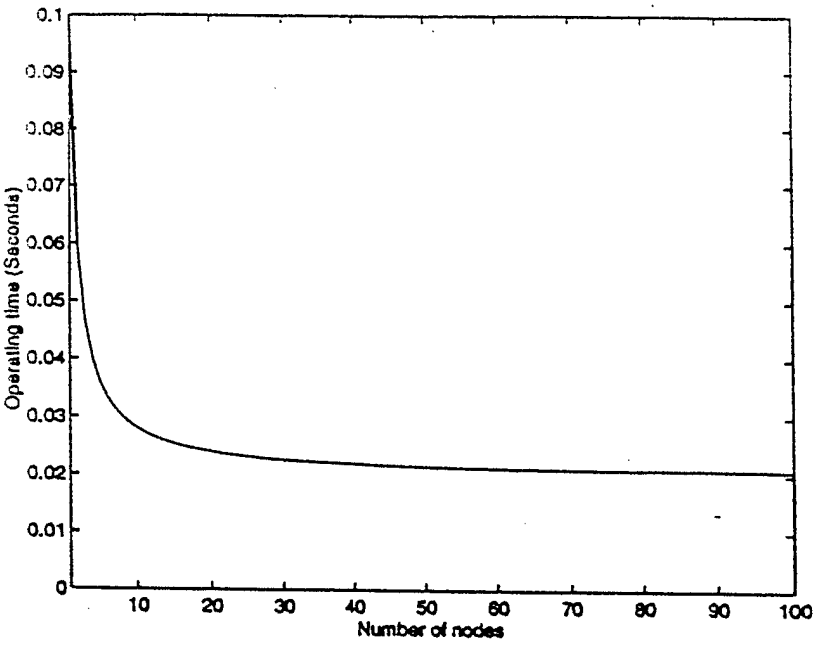


Figure 2-27
Operating time using parallel architecture with serial overhead.

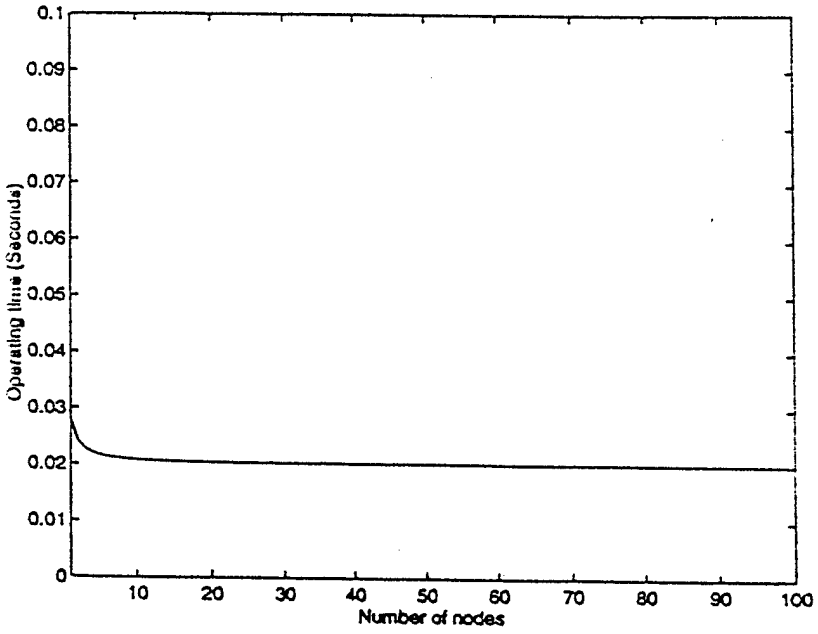


Figure 2-28
Operating time of parallel and pipeline architectures.

3.0 POTENTIAL POST APPLICATION

Over the past 20 years, commuter graphics have become the primary means of communication among technical people in an increasingly wide range of applications. CAD in engineering design, medical imaging, and visualization in scientific research are major examples. In the 1990s,

graphics have rapidly spread from technical to non-technical areas. For example, multimedia production has become very important. The use of graphics for animation is exploding into a key technology segment in the entertainment area. Also, virtual reality is on the verge of a significant role in entertainment and video games. These applications are based on low cost, real-time 3-D graphics technology.

Within this arena, POC's proposed optoelectric high speed processor system can be used for:

- Intelligent robotics
- 3-D visualization
- Medical image processing
- High definition TV (MDTV)
- 2D-to-3D conversion.

4.0 CONCLUSIONS AND RECOMMENDATIONS

4.1 Conclusions

POC's high speed processor system has achieved high performance based on innovative optical interconnections, parallel virtual buffer architecture with improvements in deep-submicron IC technology, multi-chip integration technology, and ultralow-power requirements.

First, since an optical channel can accommodate signals with >1 GHz bandwidth, many smaller bandwidth data streams can be transmitted simultaneously on a single optical channel. In previous work, POC demonstrated a channel waveguide array with a channel density of 1250 channels/cm. Second, the parallel virtual buffers scheme can increase the utilization rate of the processing elements by allocating resources dynamically around the screen as needed. Third, current VLSI technology developments allow multi-chip module packaging integrating several silicon chips on a single substrate, cramming tens of millions of transistors into a very small volume. Fourth, power consumption can be minimized, among other means, by power-down mode for the idle modules, low voltage supply and low internal voltage swing, reduced switching activities, and controlling standby leakage currents. We are employing low-power design techniques at various levels of the graphics hardware system design.

In our planned custom VLSI implementation, we take a configurable computing approach with field-programmable gate arrays (FPGAs) to prove the design concept before committing to the expensive process of full-custom silicon chip design. The configurable computing prototyping approach gives us quick feedback about our graphics processor system design. The effects of modifications and refinements can be easily evaluated to facilitate the search for an optimal solution. Our research and development has focused on the graphics architecture, guided by the HW/SW requirements and constraints. We conducted extensive performance modeling, analyzed performance bottlenecks, and now understand the hardware/software trade-off and chip design complexity. The results achieved are applicable to a variety of applications, including 3-D graphics for personal computers, high-end machines, virtual reality, animation, and other multimedia uses.

4.2 Recommendations

This Phase I project is the first step to implementation of an ultra-high speed multiprocessing system. The very high processing power is achieved by combining reconfigurable electronic processing based on fast FPGAs with high speed optical interconnections allowing nonblocking data transfer in a fully interconnected network topology.

Phase II of this project will concentrate on the further refinement of the electronic processing, especially on compact packaging issues. Reducing the size of the optoelectronic drivers to be integrated with the electronic processing core will be undertaken in the future. Efficient methods of integrating standard PCBs with optical waveguides and optical fibers will be investigated in Phase II.

The commercial goal of this project is to develop an alternative interconnection technology for IC chips that will rely on high bandwidth optical data transfer rather than on purely electronic data transfer. By integrating optical data transmission at the chip level, major bottlenecks in high density and high processing power multiprocessor systems will be opened up.

5.0 REFERENCES

1. A. Yariv, *Introduction to Optical Electronics*, 2nd Edition, McGraw-Hill, New York, NY (1976).
2. T. Gaylord and M. Moharam, "Analysis and Applications of Optical Diffraction by Gratings," Proc. IEEE 73, 840 (1985).