

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

The Design, Layout, and Characterization of VLSI Optoelectronic Chips for Free-Space Optical Interconnects

David Robert Cameron Rolston

Department of Electrical and Computer Engineering
McGill University
Montréal, Canada
July, 2000

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment
of the requirements of the degree of Doctor of Philosophy

© David Rolston, 2000



**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-69923-4

Canada

In memory of my grandfathers:

Robert N. Noakes

and

Robert C. Rolston

Abstract

The design and testing of very-large-scale-integrated optoelectronic (VLSI-OE) microchips will be described in the context of a free-space optical backplane system. The optical backplane has the potential for providing an enormous amount of bandwidth for telecommunication switching systems and massively parallel computing machines. A free-space optical backplane uses optical design techniques to relay beams of light from the surface of one microchip to the surface of another. By using light to interconnect microchips, the problems associated with high-speed electronic interconnects are avoided. By exploiting the 2-dimensional surface area of the microchips, large numbers of parallel optical interconnections are possible using minute optoelectronic devices patterned on the surface of the chips. By using appropriate optical designs and microchip layouts, massively parallel high-bandwidth interconnects can be implemented within a volume comparable with standard electronic interconnects such as buses and backplanes.

This thesis will begin by describing a specific VLSI-OE chip architecture as well as two free-space optical designs used to interconnect VLSI-OE chips. Details of the design and layout of four separate VLSI-OE chips will then be given and the results of optical and electrical testing of these chips will follow. Finally, the topic of global synchronization will then be considered. Synchronization among many VLSI-OE chips in a multiple-node system requires special attention. A novel approach of providing synchronized clock signals to a multitude of distance points will be discussed.

Résumé

La conception et le test de puces optoélectroniques-VLSI (OE-VLSI) seront décrits dans le contexte d'un système de fond de panier photonique à l'air libre. Les interconnexions optiques offrent des possibilités intéressantes pour fournir une énorme largeur de bande aux systèmes de télécommunication et aux systèmes massivement parallèles. Un fond de panier photonique à l'air libre emploie des techniques de conception optique pour transmettre par relais des faisceaux de lumière voyageant de la surface d'une puce à la surface des autres. En employant la lumière pour interconnecter des puces, les problèmes que les interconnexions électriques ont à faire face à vitesse élevée sont évités. La possibilité d'établir une interconnexion sur deux dimensions ainsi que les dispositifs optoélectroniques modelés sur la surface de la puce permettent d'offrir un grand nombre d'interconnexions optiques. De plus, la conception d'un dispositif optique approprié combiné au recours à la technologie VLSI permettent de développer des interconnexions massivement parallèles à grande largeur de bande. L'espace nécessaire pour implémenter ces interconnexions optiques est comparable à celui qu'occupent les interconnexions électriques actuelles.

Cette thèse décrira dans un premier lieu une architecture de puce spécifique à la technologie OE-VLSI. Deux conceptions optiques à l'air libre employées pour interconnecter des puces OE-VLSI seront également décrites. Des détails sur la conception et la disposition de quatre puces OE-VLSI physiquement séparées seront donnés et les résultats des tests optiques et électriques de ces puces suivront. En dernier lieu, le sujet de la synchronisation globale sera considéré. La synchronisation de plusieurs puces OE-VLSI dans un système à noeuds multiples exige une attention particulière. Un tout nouveau moyen de distribuer une horloge synchronisée aux multiples éléments d'un système sera discutée.

Acknowledgements

The guidance and friendship of my supervisor, Prof. David Plant, has been invaluable over the past 7 years, and I am grateful for the opportunity to have worked with and for him. I also owe my deepest thanks to my friend Dr. Brian Robertson, who guided me in my understanding of optics and optical experimentation. I thank my committee members, Prof. Andrew Kirk and Prof. Nicholas Rumin, they challenged me in areas that were ideal for the research I would perform. For discussions concerning phase-lock loops (among other things), I thank Prof. Gordon Roberts. And for help with clarifying my ideas regarding my patent (pending) on synchronization, I thank Mr. Thomas Adams.

Many people contributed to the optical backplane demonstrator project, each contributing to different parts of the design to varying degrees. Therefore, it is difficult to precisely define each person's contribution. However, I would like to acknowledge the following people for their major contributions to this overall research project. I sincerely apologize to anyone who I have forgot to include...

Professor David Plant had a major influence on almost every aspect of the design and construction of every demonstration system produced by this group. Prof. Andrew Kirk was part of the Phase-III design team, and contributed to the construction of Phase-III demonstration system. Prof. Ted Szymanski and Prof. H. Scott Hinton were the creators of the Hyperplane architecture that was used in this thesis as the basis for the VLSI-OE chip designs. Many of their students, such as Mr. Palash Desai, Mr. Manoj Verghesse, and Mr. Michael Kim were instrumental in the verification and simulation of the architecture. Dr. Brian Robertson, Dr. Guillaume Boisset, and Dr. Yongsheng Liu were the lead designers for the Phase-II optical design and Dr. Brian Robertson and Prof. Frank Tooley were responsible for the Phase-III optical design. Mr. Fredrick Lacroix was instrumental in the construction of the Phase-III optical system. Mr. Rajiv Iyer and Dr. Dominic Goodwill were lead designers of the Phase-II optical power supplies, and Mr. Daniel F-Brosseau was lead designer of the Phase-III optical power supply. Mr. Eric Bernier and Mr. Michael Ayliffe were responsible for the design of the optomechanical support structure for the optics and chip mounts for Phase-III. Mr. David Kabal and Mr.

Michael Ayliffe were responsible for much of the designs for the custom packaging of each VLSI-OE chip. The custom made test PCBs and flexible PCB packaging for the chips were laid-out by a collection of people. Mr. David Kabal, Mrs. Pritha Khurana, Ms. Emmanuelle Laprise, Mr. Mike Venditti, Ms. Madelaine Mony, Mr. Danny Birdie, Mr. Alan Chuah, and Mr. Feras Michael were lead designers for different PCB layouts at various times during the research. Dr. Alain Shang and Mr. Pritam Sinha were the lead designers of the charge-sense amplifiers for the Phase-III chip. Finally, discussions with Prof. Gordon Roberts and Prof. David Plant over the operation of the synchronization circuitry presented in chapter 6 were invaluable.

I would like to thank the following funding agencies that allowed me to pursue my research. The National Science and Engineering Research Council of Canada (NSERC) for their PGS-B Scholarship, the McGill-Majors/Hydro-Quebec Fellowship Office, and the Canadian Advanced Technology Association (CATA) along with La Fondation Desjardins. I would also like to thank the Canadian Institute for Telecommunications Research (CITR) for their funding of the projects I worked on and the donations of microchip fabrication from the Canadian Microelectronics Corporation (CMC).

I sincerely thank all the members of the McGill Photonics group, past and present, for their numerous technical discussions and especially their friendship. I apologize to anyone not included in this list:

Mike Ayliffe	Lukas Chrostowski	Keivan Razavi	Fred Lacroix	Tsuyoshi Yamamoto
Eric Bernier	Julien Faucher	Fred Thomas-Dupuis	Emmanuelle Laprise	Xin Xue
Danny Birdie	Amit Gupta	Eric Bisailon	Julianna Lin	Mike Venditti
Guillaume Boisset	Cosmo Girolamo	Robert Verano	Leo Lin	Pritam Sinha
Greg Brady	Wayne Hsiao	Will Phillipson	Yongsheng Liu	Alain Shang
Daniel F.-Brosseau	David Kabal	Rawiya Sherif	Tomaz Maj	Mitch Salzberg
Marc Chateauneuf	Pritha Khurana	Antoun Ghanem	Fred Mathieu	Marcos Otazo
Fan Cheng	Nam Kim	Charif Beainy	Feras Michael	Madeleine Mony

I would also like to thank my parents David and Barbara for their love, patience and guidance as well as the support and friendship of my brother Chris. And finally, to Enedina Fredes-Araya, for her love, support and encouragement as well as her comments about the flow and style of this thesis – you are the inspiration for all that I do.

Table of Contents

<u>Chapter 1: Introduction</u>	1
1.1) Optical technology	1
1.2) Electronic technology	3
1.3) The VLSI-Optoelectronic chip	6
1.4) Organization of this Thesis	7
1.5) Original Contributions	8
1.6) References	12
<u>Chapter 2: Architecture</u>	14
2.1) Introduction	14
2.2) Interconnect networks	16
2.3) The Hyperplane architecture	19
2.4) The smart pixel array	22
2.4.1) Introduction	22
2.4.2) Channel definition	23
2.4.3) SPA Protocol	24
2.4.4) SPA clocking signals	25
2.4.5) SPA logic	28
2.4.6) Data-paths of the SPA	31
2.5) A simplified SPA	32
2.6) Conclusion	35
2.7) References	36
<u>Chapter 3: Optical Interconnects</u>	37
3.1) Introduction	37
3.2) Free-Space Optical Interconnects	39
3.2.1) Introduction	39
3.2.2) Interconnect topology	41
3.2.3) Basic optical relay design	44
3.2.4) Optical modeling	45
3.2.5) Beam combination optics	49
3.2.6) Image manipulation	50
3.3) System demonstrators	51
3.3.1) Introduction	51
3.3.2) Phase-II optical system	52
3.3.3) Phase-III optical system	57
3.4) Conclusion	62
3.5) References	63

<u>Chapter 4: VLSI Optoelectronics</u>	65
4.1) Introduction	65
4.2) The Beta-Chip	66
4.2.1) Chip technology and optoelectronic specifications	66
4.2.2) Digital design	67
4.2.3) Layout	71
4.3) The Workshop-Chip	72
4.3.1) Chip technology and optoelectronic specifications	72
4.3.2) Digital design	73
4.3.3) Layout	78
4.4) The Phase-III -A Chip	79
4.4.1) Chip technology and optoelectronic specifications	80
4.4.2) Digital design	82
4.4.3) Layout	88
4.5) The Phase-III -B Chip	92
4.5.1) Chip technology and optoelectronic specifications	92
4.5.2) Digital design	93
4.5.3) Layout	97
4.6) Optoelectronics	98
4.6.1) The multiple-quantum-well	98
4.6.2) The MQW device patterning	102
4.6.3) MQW device operation	105
4.6.4) MQW device model	107
4.6.5) MQW transmitter circuit	108
4.6.6) MQW detector circuit	113
4.7) Conclusion	125
4.8) References	127
<u>Chapter 5: Experimental Results</u>	129
5.1) Introduction	129
5.2) Packaging and external control	129
5.3) Simulation and experimental results	132
5.3.1) The Beta-Chip	132
5.3.2) The Workshop-Chip	135
5.3.3) The P3A Chip	137
5.3.4) The P3B Chip	144
5.4) Conclusion	153
5.5) References	154
<u>Chapter 6: Synchronization</u>	155
6.1) Introduction	155
6.2) Standard synchronization techniques	156
6.2.1) On-chip synchronization	156

6.2.2) Board-to-board and computer synchronization	160
6.2.3) Long-distance synchronization	162
6.3) Alternative clocking structures	164
6.4) The need for a new clocking method	168
6.5) The development of the distributed synchronous clock	172
6.5.1) The target system for the distributed synchronous clock	172
6.5.2) The digital ring oscillator	174
6.5.3) The optical ring oscillator	177
6.5.4) The multiple tap-point ORO	178
6.5.5) A global clock control mechanism	180
6.5.6) Spatially separated multiple phase generation	182
6.5.7) The counter-propagating multiple pulse generator	186
6.5.8) Distributed local control	189
6.6) The distributed synchronous clock	193
6.6.1) Analytical approach to the steady-state solution	193
6.1.2) An HSpice simulation of the distributed synchronous clock	197
6.7) Conclusion	201
6.8) References	202
6.9) APPENDIX A	204
6.10) APPENDIX B	213
<u>Chapter 7: Conclusion</u>	224
7.1) Summary	224
7.2) Future Directions	227

Associated Publications

The work reported in this thesis has been published or is being published in the form of the following:

Patents:

Submitted for US and Canadian Patent (Feb 2000): "A distributed Synchronous Backplane Clocking Method", D.R. Rolston, D.V. Plant, G.W. Roberts, through the McGill Office of Technology Transfer (OTT). Patent and Trade Mark Agents: Thomas Adams & Assoc., Box 11100 Station H, Ottawa, Canada, K2H 7T8.

Refereed Journal Publications:

D.R. Rolston, B. Robertson, H.S. Hinton, and D.V. Plant, "An Optimization Technique for a Smart Pixel Interconnect using Window Clustering," *Applied Optics*, 35, no.8, pp.1220-33, March 1996.

D.R. Rolston, D.V. Plant, T.H. Szymanski, H.S. Hinton, W.H. Hsiao, M.H. Ayliffe, D. Kabal, M.B. Venditti, P. Desai, A.V. Krishnamoorthy, K.W. Goossen, J.A. Walker, B. Tseng, S.P. Hui, J.E. Cunningham, and W.J. Jan, "A Hybrid-SEED Smart pixel Array for a Four-Stage Intelligent optical Backplane Demonstrator." *IEEE Journal of Selected Topics in Quantum Electronics*, 2, no.1, pp. 97-105, Apr. 1996.

D.V. Plant, B. Robertson, H.S. Hinton, W.M. Robertson, G.C. Boisset, N.K. Kim, Y.S. Liu, M.R. Otazo, D.R. Rolston, A.Z. Shang, L. Sun, "An Optical Backplane Demonstrator System Based on FET-SEED Smart Pixel Arrays and Diffractive Lenslet Arrays," *IEEE Photon. Technol. Lett.* 7, no. 9, pp. 1057-1059 (1995).

D.V. Plant, A.Z. Shang, M.R. Otazo, D.R. Rolston, B. Robertson, and H.S. Hinton, "Design, Modeling, and Characterization of FET-SEED Smart Pixel Transceiver Arrays for Optical Backplanes," *IEEE J. Quantum. Electron.*, 32, no. 8, pp. 1391-98, Aug. 1996.

G.C. Boisset, D.R. Rolston, B. Robertson, Y.S. Liu, R. Iyer, D. Kabal, and D.V. Plant, "In situ Measurement of Misalignment Errors in Free-Space optical Interconnects", *IEEE J. of Lightwave Tech.*, 16, No. 5, May 1998.

Refereed Conference Proceedings and Technical Digests:

D.R. Rolston, D.V. Plant, H.S. Hinton, W.S. Hsiao, M.H. Ayliffe, D.N. Kabal, T.H. Szymanski, A.V. Krishnamoorthy, K.W. Goossen, J.A. Walker, B. Tseng, S.P. Hui, J.C. Cunningham, and W.Y. Jan, "Design and Testing of a Smart Pixel Array for a Four-Stage Optical Demonstrator", *IEEE LEOS Topical Meeting on Smart Pixels*, pp. 30-31 (1996)

D.V. Plant, B. Robertson, H.S. Hinton, M.H. Ayliffe, G.C. Boisset, D.J. Goodwill, D. Kabal, R. Iyer, Y.S. Liu, D.R. Rolston, M. Venditti, T.H. Szymanski, W.M. Robertson, M.R. Taghizadeh, "Optical, Optomechanical, and Optoelectronic Design and Operational Testing of a Multi-Stage Optical Backplane Demonstrator System", Proceedings of MPPOI '96, IEEE Computer Society, Maui, Hawaii, pp. 306-312, Oct. 27-29, 1996.

K.E. Davenport, H.S. Hinton, D.J. Goodwill, D.V. Plant, D.R. Rolston, and W.S. Hsiao, "A Hyperplane Smart Pixel Array for Packet Based Switching", IEEE LEOS Topical Meeting on Smart Pixels, pp. 32-33 (1996)

D.N. Kabal, G.C. Boisset, D.R. Rolston and D.V. Plant, "Packing of Two-Dimensional smart pixel arrays", IEEE LEOS Topical Meeting on Smart Pixels, pp. 53-54 (1996)

D.V. Plant, B. Robertson, H.S. Hinton, W.M. Robertson, G.C. Boisset, N.K. Kim, Y.S. Liu, M.R. Otazo, D.R. Rolston, A.Z. Shang, "Optical Backplane Demonstrators Based on FET-SEED Smart Pixel Arrays," Proceedings of the IEEE LEOS Annual Meeting, pp. 238-239 (1994).

B. Robertson, G.C. Boisset, H.S. Hinton, Y.S. Liu, N.H. Kim, M.R. Otazo, D. Pavlasek, D.V. Plant, and D.R. Rolston, "Design of a Lenslet Array Based Free-Space Optical Backplane Demonstrator," ICO Optical Computing Conference 1994, Institute of Physics Conference Series Number 139, pp. 223-226 (1994).

D.V. Plant, B. Robertson, H.S. Hinton, W.M. Robertson, G.C. Boisset, N.H. Kim, Y.S. Liu, M.R. Otazo, D.R. Rolston, A.Z. Shang, L. Sun, "A FET-SEED Based Optical Backplane Demonstrator," ICO Optical Computing Conference 1994, Institute of Physics Conference Series Number 139, pp. 145-148 (1994).

D.V. Plant, B. Robertson, H.S. Hinton, G.C. Boisset, N.H. Kim, Y.S. Liu, M.R. Otazo, D.R. Rolston, A.Z. Shang, and W.M. Robertson, "Micro-Channel Based Optical Backplane Demonstrators Using FET-SEED Smart Pixel Arrays," Proceedings of the SPIE, 2400, pp. 170-174 (1995).

D.R. Rolston, B.R. Robertson, D.V. Plant, and H.S. Hinton, "Design Space Analysis of a Lenslet Based Optical Relay System Interconnecting Smart Pixel Arrays," Technical digest of the OSA Topical Meeting on Optical Computing, pp. 102-104 (1995).

D.V. Plant, B. Robertson, G.C. Boisset, N.K. Kim, Y.S. Liu, R.M. Otazo, D.R. Rolston, A.Z. Shang, H.S. Hinton, W.M. Robertson, "16-Channel FET-SEED Based Optical Backplane Interconnection," Technical digest of the OSA Topical Meeting on Optical Computing, pp. 272-275 (1995).

D.V. Plant, B. Robertson, H.S. Hinton, M.H. Ayliffe, D.R. Rolston, et al. "A multistage CMOS-SEED optical backplane demonstration system." 1996 International

Topical Meeting on Optical Computing. Technical Digest. - (OC96, Sendai, Japan, 21-25 April 1996). Tokyo, Japan: Japan Soc. Appl. Phys, Vol. 1, pp.14-15, 1996

D.V. Plant, B. Robertson, H.S. Hinton, W.M. Robertson, D.R. Rolston, et al. "Optical backplane demonstrators using micro-optics and smart-pixel transceiver arrays." CLEO '95. Summaries of Papers Presented at the Conference on Lasers, Washington DC, USA: Opt. Soc. America, pp. 396-7, 1995

D.V. Plant, B. Robertson, H.S. Hinton, W.M. Robertson, D.R. Rolston, et al., "A FET-SEED smart pixel based optical backplane demonstrator." Optical Computing - Proceedings of the International Conference. (Optical Computing, Edinburgh, UK, 22-25 Aug. 1994). Edited by: Wherrett, B.S.; Chavel, P. Bristol, UK: IOP Publishing, pp. 145-8, 1995.

B. Robertson, G.C. Boisset, H.S. Hinton, Y.S. Liu, D.R. Rolston, et al., "Design of a lenslet array based free-space optical backplane demonstrator." Optical Computing - Proceedings of the International Conference. (Optical Computing, Edinburgh, UK, 22-25 Aug. 1994). Edited by: Wherrett, B.S.; Chavel, P. Bristol, UK: IOP Publishing, pp. 223-6, 1995.

D.V. Plant, B. Robertson, H.S. Hinton, G.C. Boisset, D.R. Rolston, et al. "Microchannel based optical backplane demonstrators using FET-SEED smart pixel arrays." Proceedings of the SPIE - The International Society for Optical Engineering (Optoelectronic Interconnects III, San Jose, CA, USA, 8-9 Feb. 1995), Vol. 2400, pp. 170-4, 1995.

D.V. Plant, B. Robertson, H.S. Hinton, W.M. Robertson, D.R. Rolston, et al. "Optical backplane demonstrators based on FET-SEED smart pixel arrays." LEOS '94. Conference Proceedings. IEEE Lasers and Electro-Optics Society 1994 7th Annual Meeting (LEOS '94 Boston, MA, USA, 31 Oct.-3 Nov. 1994) vol.1, pp. 238-9, 1994.

Chapter 1: Introduction

1.1) Optical technology

Switching systems are quickly approaching the limit of metal trace-line technology just as long-distance communications based on copper began to reach its limit over three decades ago when fiber-optics was first proposed as an alternative [ref 1]. The relatively low bandwidth and attenuation of copper wire limited the data rate that could be communicated over long distances. The fiber-optic alternative continued to grow in complexity due to numerous technological advances making it more affordable and easier to implement. Multi-mode fibers gave way to single-mode fibers that allowed for higher bit-rates. Lower dispersion and attenuation in fibers allowed longer distances to be achieved before re-amplification of the signal was required. Further advances, such as the “all-optical” amplification of Erbium-doped fiber (an “in-fiber” Laser gain medium) [ref 2], allowed an optical signal to be amplified without first being converted into an electrical signal. This lowered the overall cost of fiber-based systems and increased the reliability. Another advance in fiber-based systems was the use of multiple wavelengths in a single fiber; called wavelength division multiplexing (WDM) [ref 3]. Different wavelengths of light could be individually modulated and introduced into the same fiber. At the receiving-end of the fiber, these wavelengths could be split apart allowing for several “virtual” fibers in one. Dense-WDM (DWDM) technology [ref 4] has allowed for possibly hundred’s of wavelengths to be merged into one fiber and advances in Erbium-doped amplifiers have

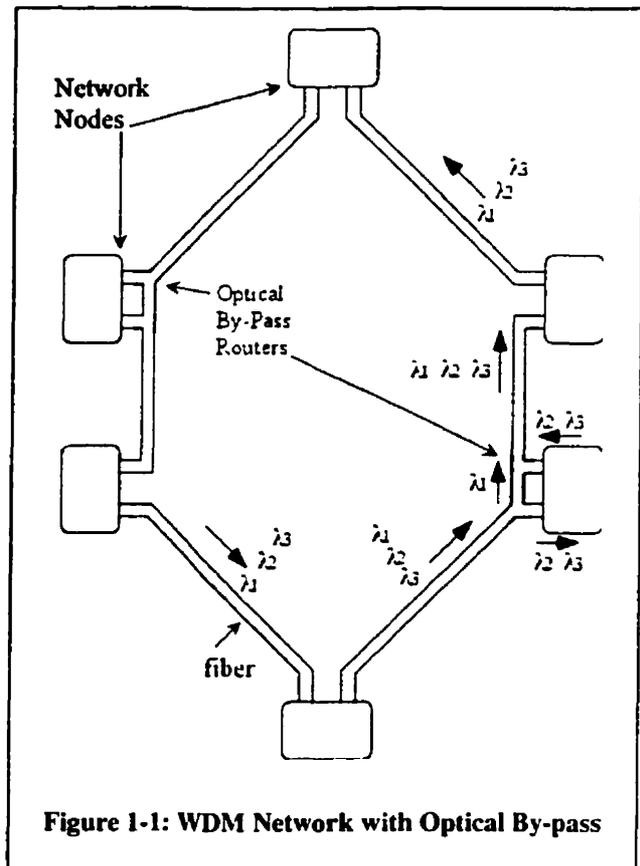


Figure 1-1: WDM Network with Optical By-pass

allowed DWDM to maintain signal integrity and bandwidth for long distance applications. The next trend, and the ultimate in fiber-based systems, will be optical-flow-switching (OFS) [ref 5]. Optical-flow-switching would allow DWDM fibers to dynamically switch wavelengths in an “all-optical” manner. At the present time, DWDM fibers must use electronic hardware to interrogate asynchronous-transfer-mode (ATM) or synchronous-optical-network (SONET) data to determine if it is destined for a particular node [figure 1-1]. This conversion is a very noticeable bottleneck in the network. A rudimentary form of OFS called “optical by-pass” already exists in some forms, which allows certain wavelengths to pass nodes transparently as long as none of the data contained in that wavelength is ever destined for that node. However, the network loading must already be known so that a specific topology can be imbedded into the physical layer. An optical-flow-switched network would allow for a tremendous advantage over simple optical by-pass because the interrogation could be carried-out without the need for electronic hardware.

Although OFS would be a revolution in telecommunications, it is exceedingly difficult to envision the type of technology that this would entail - barring a major breakthrough in photonic device research. It is even more difficult to envision an available cost-effective “all-optical” solution. An OFS would have to interrogate hundreds of gigabit per second channels carrying sequential pulses of light and then re-direct all the appropriate data to the correct nodes in an “all-optical” manner. However, for the foreseeable future, the electronic integrated circuit is the only technology that can provide the means for interrogating and controlling data. The transistor density and the speed of

International Technology Roadmap for Semiconductors

Chip Parameter	Year						
	1997	1999	2002	2005	2008	2011	2014
Technology (nm)	250	180	130	100	70	50	35
Number of Package Pins	1136	1400	1915	2619	3581	4898	6700
On-Chip Frequency (MHz)	750	1250	2100	3500	6000	10000	16903
Chip-to-board (off-chip) peripheral buses (MHz)	250	480	885	1035	1285	1540	1878

<http://www.itrs.net/itrs/publntsr.nsf> (Feb. 24th, 2000)

Figure 1-2: SIA Roadmap for Semiconductors

typical electronic circuits have been steadily doubling every two years and they are projected to continue doubling for some time [figure 1-2]. Therefore, a complete “all-optical” switching network may be not be immediately required, especially if high-speed electronics can be more closely combined with the optical systems. This thesis will explore some of the chip design techniques used to bring high-speed processing closer to the optical layer to create a hybrid between electronics and optical switching networks.

1.2) Electronic technology

In the same way that electronics may help optical systems achieve high switching performance, optics may help electronics increase its performance as well. The same issues that led to the decline in the use of metal conductors for long-haul communications are now starting to affect systems at the backplane, bus, printed-circuit-board

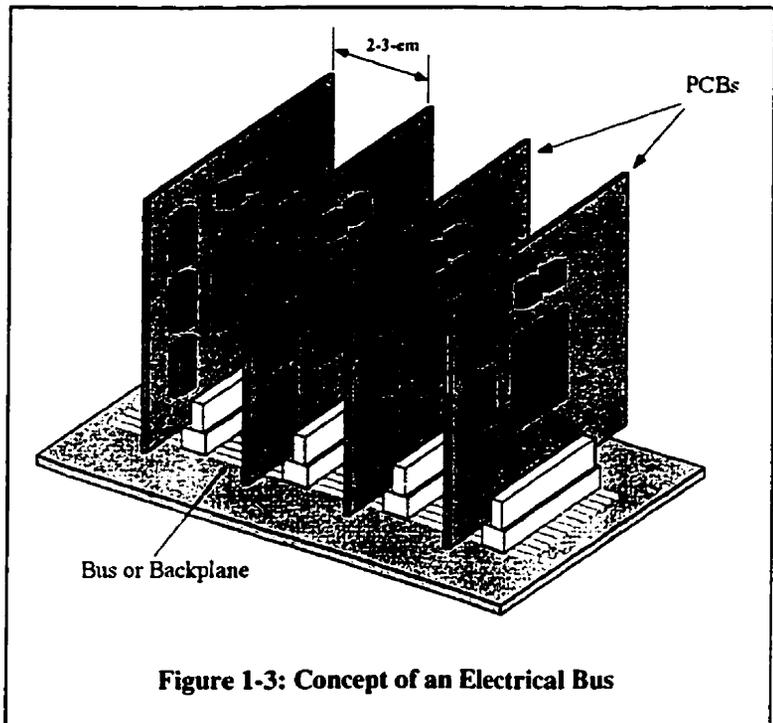


Figure 1-3: Concept of an Electrical Bus

(PCB), and even chip level [figure 1-3] due to high data rates. The exceedingly high data

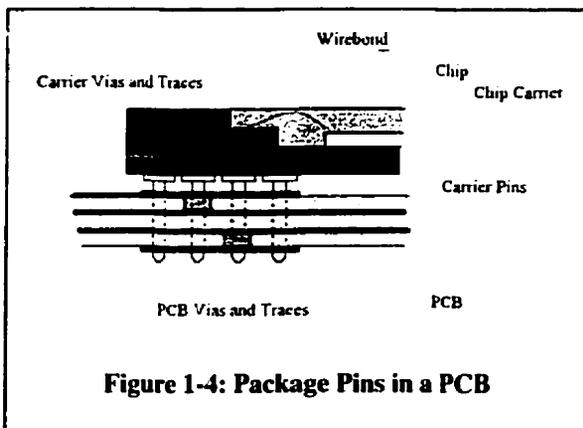


Figure 1-4: Package Pins in a PCB

rates and the large numbers of electrical data paths available from high-end microprocessors are beginning to cause difficulty with electronic communications from board-to-board and chip-to-chip. The transmission line effects of small lengths of wire are more noticeable when operated at high data

rates. Poorly terminated buses and unwanted reflections from pins and connectors also compromise signal integrity [figure 1-4][figure 1-5]. The high power-consumption from

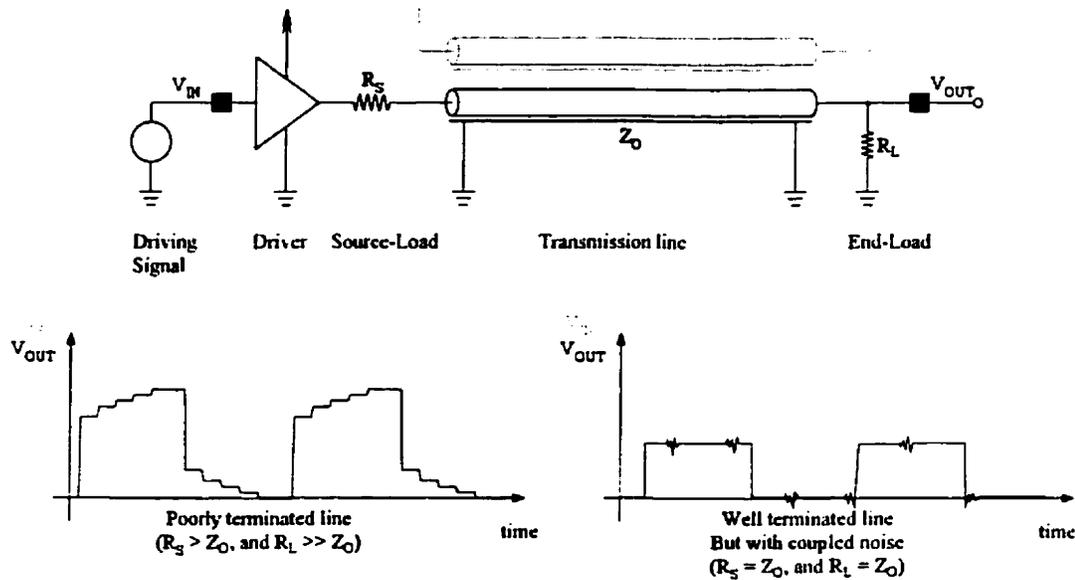


Figure 1-5: Examples of Electrical Transmission Line Responses

transceiver drivers on terminated buses can also cause thermal problems, which can be very difficult to design for, especially when very high data rates are used. Multiple signals changing simultaneously can also cause power and ground bounce as well as induce switching noise proportional to the rate of change of the signal, sometimes called “ $\delta V/\delta t$ ” noise or “ ΔI ” noise [ref 6,7].

The Sun Microsystems UltraSPARC-IIi CPU module is an example of a situation where the on-chip speed of the CPU exceeds 300-MHz and it has a 64-bit wide data bus to external secondary

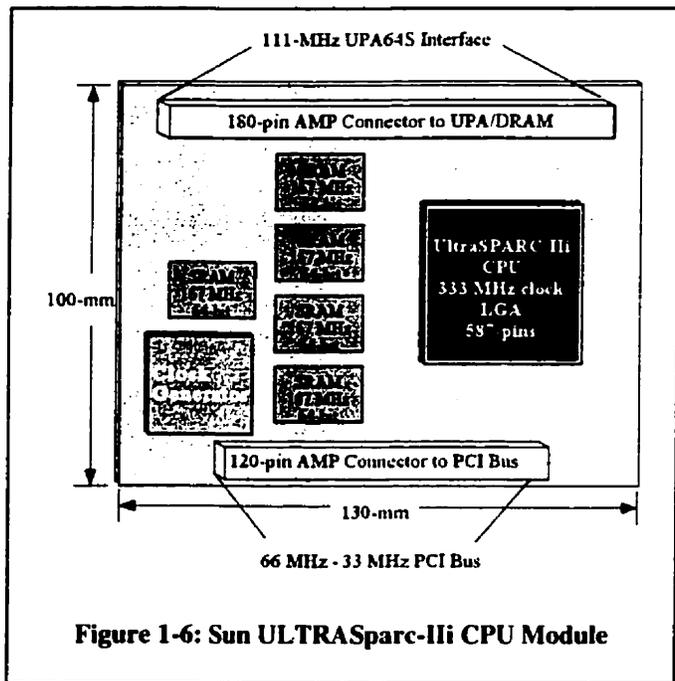


Figure 1-6: Sun ULTRASparc-IIi CPU Module

cache. However, due to transmission line effects, the CPU can access external memory at only 167-MHz (this is a dedicated CPU to memory bus) [figure 1-6] [ref 8]. The data rate off and on the module is again limited by the PCI bus to 66-MHz [ref 9]. Another method used to increase speed by eliminating transmission line effects has been the use of multi-chip-module (MCM) and system-on-a-chip designs to compact as much processing, memory and control as possible onto one package or chip, respectively [ref 10,11].

Although MCM technology allows higher speeds within the packages, because

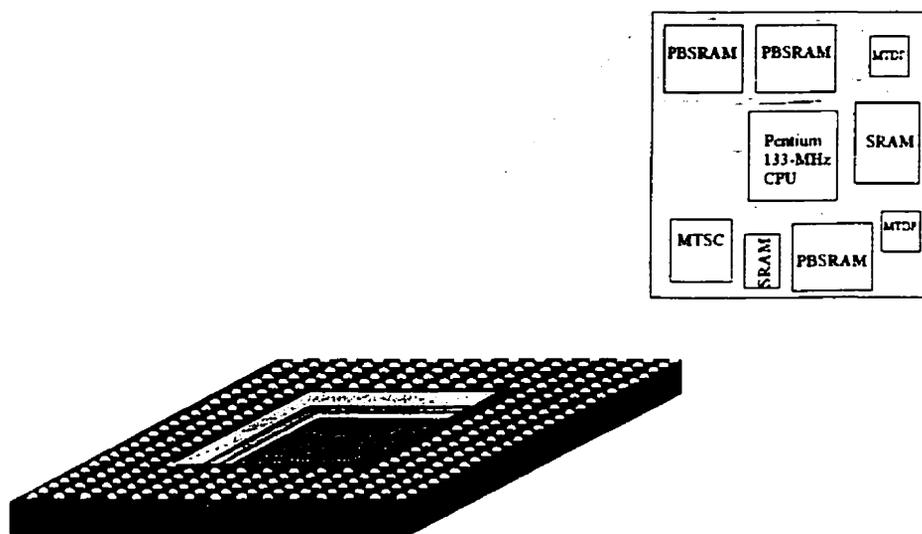


Figure 1-7: Example of a Pentium CPU in an MCM Ball-Grid-Array Package

the distances and RC-effects are kept to a minimum, the off-chip speeds are limited to below 100-MHz. An example of this technology is the MCM-D package used to interconnect a Pentium-II with multiple cache memories on a 4-layer deposited substrate within the 32-mm x 32-mm cavity MCM [figure 1-7] [ref 12]. Although the internal speed was over 200-MHz, the external data rates were limited by the bus (see Chapter 6 – Synchronization).

The advantages of using well-established silicon processing to employ the decision and comparison part of the switch are numerous. The relative ease with which one can implement complicated architectures, coupled with the fabrication infrastructure of silicon electronics, makes silicon IC technology a compelling choice for any switching system. However, as the limitations of electronic transmission lines are reached for

backplane level systems, optical technology can be judiciously applied. If both the highly integrated silicon processing technology and the interconnection of optical data paths are merged at the appropriate interface, both optical networks and computational systems will benefit.

1.3) The VLSI-Optoelectronic chip

Many of the bandwidth, noise, and thermal problems associated with standard electronic communication technologies may be alleviated by using the 3-dimensional connectivity of very-large-scale-integrated optoelectronic (VLSI-OE) chips and optical

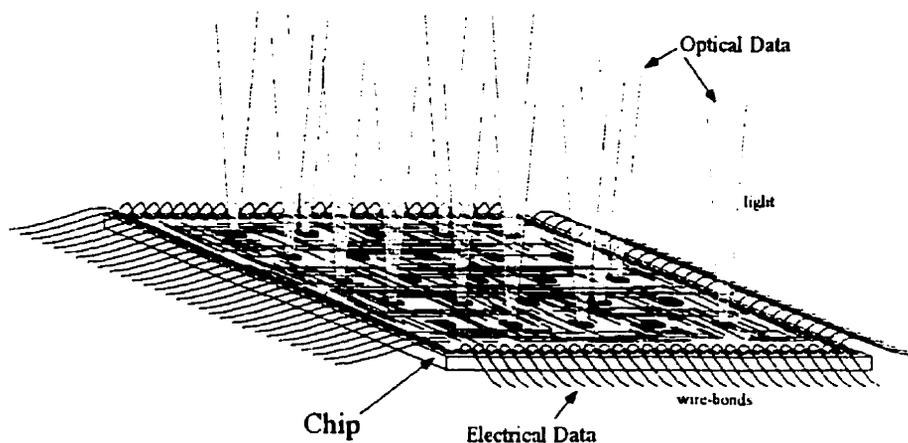


Figure 1-8: Concept of a VLSI-OE Chip

interconnects [ref 13,14,15,16]. Optoelectronic devices connected to the surface of silicon processing chips will allow data to move on and off chips at very high rates and with fractions of the power required to operate electronic bus systems [figure 1-8].

The type of optical interconnect is somewhat immaterial with respect to the design of the VLSI-OE chip. The basic structure, layout and testing of the VLSI-OE chip does not drastically change when either free-space or fiber based interconnects are used to connect between chips. However, the assumption that the basic "form-factor" of a switching system will be comprised of multiple printed-circuit boards in a card-rack or chassis indicates that a micro-optic free-space optical design may be more desirable than a fiber-based system. If the required density of optical interconnects between chips or

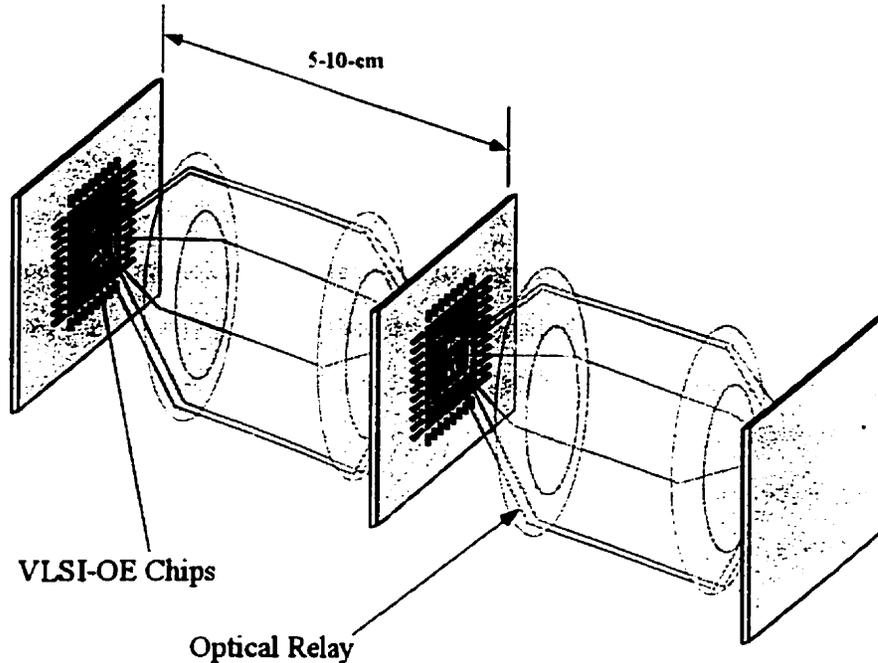


Figure 1-9: Concept of a Free-Space Optical Interconnect

boards must continually double every two years, then achieving thousands of individual optical fiber connections between boards separated by only a few centimeters may be a significant challenge to fiber connector technology. The free-space alternative may allow many thousands of optical beams [figure 1-9] to be relayed from chip-to-chip as long as alignment and optomechanical structures are sufficiently robust.

1.4) Organization of this Thesis

Chapter 2, the Architecture, will introduce a network topology suitable for either switching or computing needs. The architecture is based on a reconfigurable crossbar interconnect that allows both optical and electrical forms of data communication to interact at the chip level. Chapter 3, the Optical Interconnect, describes a representative set of demonstration systems that were used as the physical interconnect layer for the proposed architecture. It will demonstrate some of the methods used to optically interconnect optoelectronic chips. The next chapter, VLSI-Optoelectronics, describes in detail 4 chip iterations based on the architecture presented earlier. The complexity of each chip was increased from that of the previous design. The impact of the optical design on

chip layout as well as the impact standard chip layout practices had on the optical design will be highlighted. In chapter 5, Experimental Results from tests performed on each chip will be provided. Quantitative data on chip performance is highlighted. This chapter also highlights some of the issues that brought about design revisions between iterations. Chapter 6 deals with some of the design issues in the synchronization of numerous printed-circuit-boards operating at high data rates. This synchronization technique may also be used in systems other than the optical interconnect.

1.5) Original Contributions

When this work began in 1994, there were only a handful of researchers and companies attempting to build free-space optical interconnects. A group within AT&T Bell Labs had begun work on optical Banyan networks that led to a series of system demonstrators [ref 17]. These demonstrators were all-optical cross-connects that manipulated optical data by converting the information to electrical data and then reconverting it back to optical data. The European Project, ESPRIT, was also investigating free-space optical interconnects to alleviate the bandwidth bottleneck at the chip level [ref 18]. The company NEC created novel optical interconnect topologies [ref 19], but required optical fiber to supply the data streams. What none of the groups at that time had done was to demonstrate a method of fully integrating the on and off chip electrical bandwidth of VLSI CMOS silicon processing with optoelectronic devices and free-space optics for data routing. The work presented in this thesis addressed this specific issue.

In this thesis, one of the first hardware implementations of a CMOS microchip for an optical backplane switching architecture is presented. The chip demonstrated the ability to handle both high-speed electrical data and high-speed optical data. The chip was able to convert electrical data to optical data, optical data to electrical data, and optical data back to optical data. These processes were enhanced by incorporating a method to dynamically change between states depending on the type of data present. Three more VLSI-OE chips followed, where each considered more complicated aspects

of the optical backplane design. Although the same basic architecture was still used, several aspects dealing with implementation had to be addressed.

The objective of the project was to improve upon existing electronic backplane technology, while adhering design requirements, such as a small physical size for the optical backplane and a 2-cm pitch of 6-U PCBs. To achieve this, one of the most significant design parameters became the optical signal density. A channel density of greater than 1000 optical channels per cm^2 was set as the goal. This required that each VLSI-OE chip have a very dense array of optoelectronic devices over its surface.

The ideal placement of optoelectronic devices on the surface of a VLSI-OE chip was a major issue that impacted the design of the optical system as well as the layout of the microchip; a suitable compromise had to be found. Since the architecture was based on a matrix of repeatable cells (called Smart Pixels), the most ideal VLSI layout technique was to have each cell contain an optical receiver element and an optical transmitter element. This would entail a regular array of transmitter/receiver pairs across the surface of the chip, but would lead to a more complicated optical design. However, an optical design, which grouped the optical receivers on one side of the chip the optical transmitters on the other side of the chip would allow a more simple optical system. Unfortunately, it would require thousands of signal trace lines to be routed into dense arrays of optoelectronic devices on the chip. The solution was to find a compromise between the two requirements.

As a compromise between a regular array of receiver/transmitter pairs, and unique receiver and transmitter groups, the method of window clustering was employed. Two of the chips described in this thesis were the first demonstrations of clustered-array optoelectronic chips. The clustering technique analyzed in a previous work [ref 20], provided a means to obtain the desired channel density while allowing flexibility in both the optical and electronic designs. The layout technique for logic that corresponded to the clustered-array of optoelectronic devices was also a novel contribution to the study of VLSI-OE chips. It demonstrated several techniques that relied on symmetric layout and cell repeatability in order to reduce the overall complexity of the layout.

Novel VLSI-OE testing methods were also a large part of the contributions of this work. The simultaneous transfer of high-speed electrical data to the VLSI-OE chip

combined with both the generation of optical input data and the detection of optical output data involved not only complicated optical test setups, but also the associated electrical control of the chip.

Finally, due to an architecture that required several clocks to be distributed among printed circuit boards in the optical backplane, a novel circuit called the Distributed Synchronous Clock (DSC) was developed. Although intended for the architecture present herein, the DSC could be placed into any system that requires a simultaneous triggering of many physically separated points. The physical medium among nodes can often cause unwanted skew for trigger signals, circuits such as H-tree distribution networks have been used to equalize delay through a network. Occasionally more circuit-design is required; in terms of phase-locking circuits, memory buffering, and re-clocking of data. Effects such as time-degradation of components and changing thermal gradients also cause dynamic variation in the exact timing of triggering signals. However, the proposed DSC circuitry not only allows multiple points to be synchronized together, but it also allows dynamic compensation when changes in the system occur. This circuit may allow massively parallel architectures to improve performance by eliminating the circuitry overhead required for skewed clocks. It may even provide a means of synchronizing earth-orbiting satellite networks.

Summary of Novel Contributions:

- 1) The first completely custom hardware implementation of the Hyperplane smart pixel architecture.
- 2) The characterization of one of the first fully electrical-in/optical-out, optical-in/electrical-out, optical-in/optical-out VLSI-OE chips.
- 3) The development of strategies for optimal VLSI layout techniques with respect to optical system integration. This includes the first fabricated clustered optoelectronic device array on a VLSI chip.
- 4) Unique testing procedures for VLSI-OE chips including the design and implementation of a microscope test-bench for on-chip probing of silicon chips before optoelectronic attachments were made.

5) The novel creation of a means of synchronizing multiple chips or PCBs and dynamically correcting for clock-edge misalignments using phase-lock loop techniques.

1.6) References

- [1] K.C. Kao, G. Hockham, IEE Proceedings, Vol. 113, 1966, p. 1151
- [2] M. X. Ma, et al., "765 Gb/s over 2,000 km Transmission Using C- and L-Band Erbium Doped Fiber Amplifiers", Optical Fiber Communication Conference, International Conference on Integrated Optics and Optical Fiber Communication (OFC/IOOC '99) Technical Digest, 1999, pp. PD16/1-PD16/3 Suppl.
- [3] C. Scheerer, C. Glingener, "WDM and ETDM for future optical transmission systems", IEEE Global Telecommunications Conference (GLOBECOM 1998), Vol. 2, 1998, pp. 1007 -1011
- [4] M. Fukutoku, N. Shibata, "16 WDM optical packet routing experiment over 640-km transmission distance at a data rate of 2.5-Gb/s", Optical Fiber Communication Conference (OFC/IOOC '99) Technical Digest, Vol. 3, 1999, pp. 168 -170
- [5] E. Modiano, A. Narula-Tam, "Mechanisms for Providing Optical Bypass in WDM-based Networks", Optical Networks, Vol. 1, No. 1, Jan 2000, pp. 10-20
- [6] B.C. Martin, "A 2.2-ns 2.5/3.3-volt BiCMOS bus transceiver using pass-NMOS BiCMOS design", Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting, 1996, pp. 97 -100
- [7] Y.I. Ismail, E.G. Friedman, J.L. Neves, "Power dissipated by CMOS gates lossless transmission lines", Proceedings of the International Symposium on Low Power Electronics and Design, 1998, pp. 139 -141
- [8] UltraSPARC-III CPU Module, Sun Microsystems Data Sheets
- [9] G. C. L. Boisset, Optomechanics and Optical Packaging for Free-Space Optical Interconnects, Ph.D. Thesis, McGill University, Montreal, Canada, 1998.
- [10] T. Isshiki, P. Garay, J. Ramirez, V. Maheshwari, W.W-M. Dai, "A silicon-on-silicon field programmable multichip module (FPMCM) integrating FPGA and MCM technologies", IEEE Transactions on Advanced Packaging: Components, Packaging, and Manufacturing Technology, Part B, Vol. 18, No. 4, Nov. 1995, pp. 601 -608
- [11] A.M. Rincon, G. Cherichetti, J.A. Monzel, D.R. Stauffer, M.T. Trick, "Core design and system-on-a-chip integration", IEEE Design & Test of Computers, Vol. 14, No. 4, Oct.-Dec. 1997, pp. 26 -35
- [12] E. Hirt, M. Scheffler, J.-P. Wyss, "Area I/O's potential for future processor systems", IEEE Micro Vol. 18, No. 4, July-Aug. 1998, pp. 42 -49
- [13] A.V. Krishnamoorthy, J.E. Ford, F.E. Kiamilev, R.G. Rozier, S. Hunsche, K.W. Goossen, B. Tseng, J.A. Walker, J.E. Cunningham, W.Y. Jan, M.C. Nuss, "The AMOEBA switch: an optoelectronic switch for multiprocessor networking using dense-WDM", IEEE Selected Topics in Quantum Electronics, Vol. 5, No. 2, March-April 1999, pp. 261 -275
- [14] F.E. Kiamilev, A.V. Krishnamoorthy, "A high-speed 32-channel CMOS VCSEL driver with built-in self test and clock generation circuitry", IEEE Selected Topics in Quantum Electronics, Vol. 5, No. 2, March-April 1999, pp. 287 -295
- [15] J. Shibata, T. Kajiwara, "Optics and electronics are living together", IEEE Spectrum Vol. 26, No. 2, Feb. 1989, pp. 34 -38

- [16] A.J. Moseley, M.Q. Kearley, R.C. Morris, D.J. Robbins, J. Thompson, M.J. Goodwin, "Uniform 8x8 array InGaAs/InP multiquantum well asymmetric Fabry-Perot modulators for flip-chip colder bond hybrid optical interconnect", *Electronics Letters*, Vol. 28, No. 1, Jan. 1992, pp. 12-14
- [17] F.B. McCormick, T.J. Cloonan, A.L. Lentine, J.M. Sasian, R.L. Morrison, M.G. Beckman, S.L. Walker, M.J. Wojcik, S.J. Hinterlong, R.J. Crisci, R.A. Novotny, H.S. Hinton, "Five-stage free-space optical switching network with field-effect transistor self-electro-optic-effect devices", *Applied Optics*, Vol. 33, 1993, pp. 5153-5171.
- [18] J. W. Parker, "Optical Interconnection for Advanced Processor Systems: A Review of the ESPRIT II OLIVES Program", *J. of Lightwave Technology Lett.*, Vol. 9, No. 12, December 1991, pp. 1764-1772.
- [19] S. Araki, M. Kajita, K. Kasahara, K. Kubota, K. Kurihara, I. Redmond, E. Schenfeld, T. Suzuki, "Experimental free-space optical network for Massively Parallel Computers", *Applied Optics*, vol.35, no.8, 10 March 1996, pp.1269-81.
- [20] D.R. Rolston, B. Robertson, H.S. Hinton, and D.V. Plant, "An Optimization Technique for a Smart Pixel Interconnect using Window Clustering," *Applied Optics*, 35, no.8, pp.1220-33, March 1996.

Chapter 2: Architecture

2.1) Introduction

The system architecture presented in this chapter provides an alternative to the switching architectures currently employed in telecommunication switching nodes. A telecommunications switching network is typically composed of many nodes each linked to one another with high-speed serial links in a web-like structure. Each node is responsible for redirecting messages to other nodes in the network so they can eventually

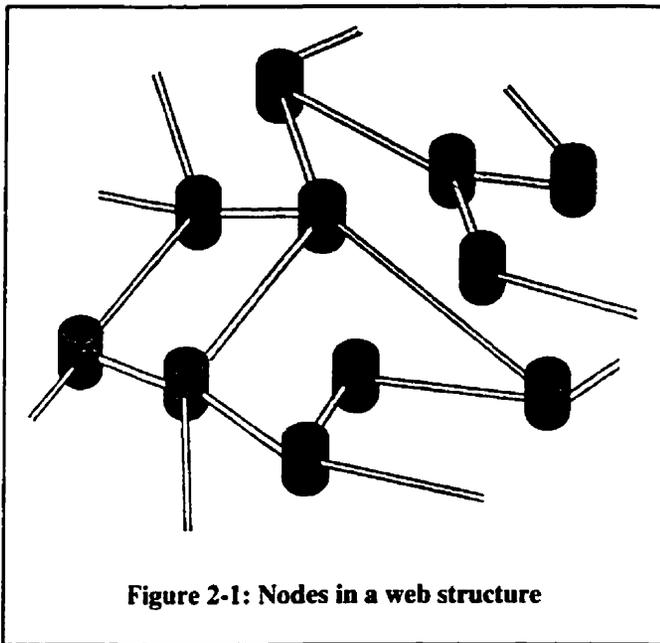


Figure 2-1: Nodes in a web structure

be routed to their proper destinations [figure 2-1].

Typically, the links between nodes are high bandwidth optical fibers transmitting up to a terabit per second of information [ref 1].

Transmission and encoding techniques, such as wavelength division multiplexing (WDM) [ref 2], allow multiple gigabit per second channels to be manipulated

so that congestion or failures in the

network can be avoided. However, these high-speed channels must eventually be decomposed into many parallel sets of slower speed data that must be processed by electronic computing hardware.

Although the speed at which calls can be serviced has dramatically increased over the last 50 years, the fundamental method of switching has not. Most systems use relatively old concepts that are continuously adapted to faster integrated technology. Just as early telephone exchanges required a manual patch between one caller and another, ultimately the switching fabric is responsible for connecting two points together. A typical switching fabric is the Banyan $N \times N$ cross-connect. The Banyan interconnect is composed of multiple layers of two by two switching elements and uses multiple cross-over connections to map N -inputs to N -outputs [ref 3]. The switching matrix uses a

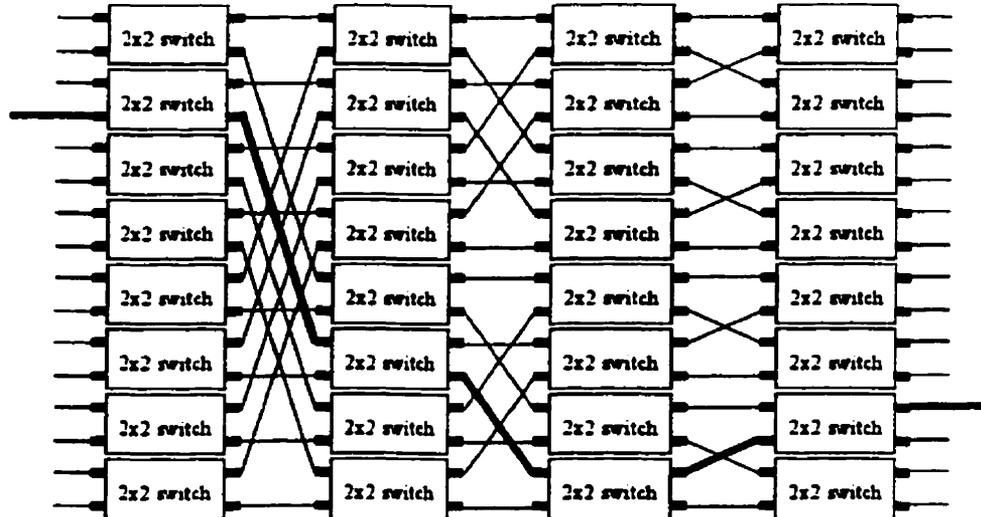


Figure 2-2: A 2-D Banyan switching fabric

butterfly-interconnect pattern. The Banyan cross-connect is a non-blocking switch since it is possible for any input to connect to any output [figure 2-2]. However, there can be contention between two inputs and a single output, which causes one of the inputs to be “dropped”.

The Banyan interconnect is one particular hardware structure used in telecommunication networks and was first introduced for circuit-switched interconnects. A circuit-switched network creates a fixed path through all the required nodes of the network from input to output for the duration of the transaction. Each switch in the network is set in one of two states that enable an input to communicate with an output. There are many variations of the Banyan interconnect, each try to limit the amount of resources and minimize cost, while at the same time maximize bandwidth and minimize the message-loss probability. However, these networks can at times take a remarkably long time to transfer data or place a call. The latency associated with most telecommunication networks is primarily due to the limited resources at the switch that routes the data. Since many inputs may be competing for the same output, a network usually requires large amounts of memory buffering at each node, this causes latency in the communication. One of the only switches that is non-blocking, and has no contention is the fully connected crossbar interconnect. Theoretically, it can connect all points of a

system simultaneously. Unfortunately, it also requires a prohibitively high number of interconnections, which is not cost effective for all but the smallest systems (less than a few 10's of nodes).

The system architecture presented in this chapter offers a possible method to employ the crossbar interconnect and alleviate some of the problems with traditional switching systems. The architecture presented in the following sections is a specific implementation of the more generic *Hyperplane* switching architecture [ref 4]. The evolution of optoelectronic and microelectronic technology is an essential part in the development of this architecture. The ability to fabricate suitable optoelectronic devices that can convert between electronic data and optical data is key in the development of this architecture. The ability to integrate optoelectronic devices with standard microelectronic processing circuitry, such as CMOS VLSI, is another crucial factor in the development of this architecture. By using each of these elements, a method which takes advantage of the inherent parallel nature of optical imaging can be constructed and allows a version of the crossbar interconnect to be implemented using a reasonable amount of hardware.

In this chapter, a general outline of typical interconnection networks will be given along with two examples of optically interconnected systems. The specific interconnect architecture will then be discussed along with all electronic sub-components required to implement the interconnect. A device called a *Smart Pixel Array* will be presented as the core of the architecture and a high-level description of its operation will be given. Several definitions, such as the number of optical channels and the definition of an optical packet will be provided as well. The signals required to control and synchronize the entire system are also shown. In this chapter, optical channels are modeled as ideal point-to-point connections between smart pixel arrays.

2.2) Interconnect networks

The crossbar interconnect has the most complete coverage of virtually any network. A fully connected crossbar will allow any input to connect to any output at any time. If the crossbar has N-inputs, there must be N-groups of N-outputs to allow full coverage.

A fully connected crossbar requires a tremendous amount of fan-out although virtually no switching is required [figure 2-3]. Although the crossbar may be highly desirable in terms of overall coverage, it cannot be implemented due to the large number of interconnect lines required for even the most modest size system. Other

interconnection embeddings such as the Mesh, the HyperCube and the Banyan networks each use some form of switching in order to compensate for a fewer number of point-to-point connections, but are far easier to implement [ref 5,6]. For example, the 3-D Mesh network has six connections for each node [figure 2-4]; thus each node requires only six inputs and six outputs. A message issued by one node in the

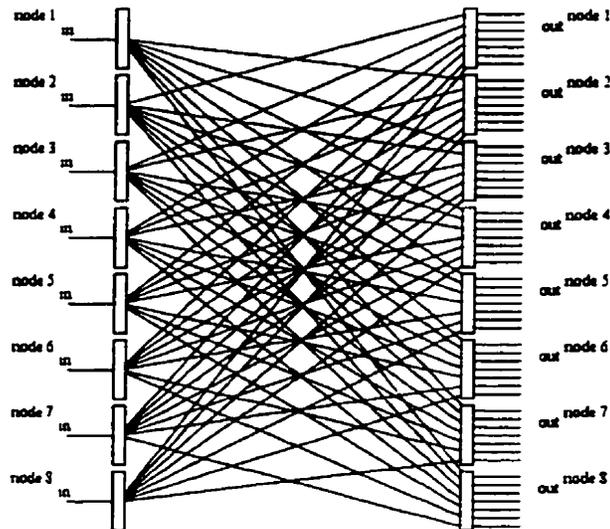


Figure 2-3: A fully-connected crossbar

Mesh must find its way through many adjacent nodes in order to arrive at the proper destination. This requires that each node in the network have the ability to redirect incoming messages towards nodes that are closer to the appropriate destinations. Unlike the fully connected crossbar, embeddings that involve switching require some form of

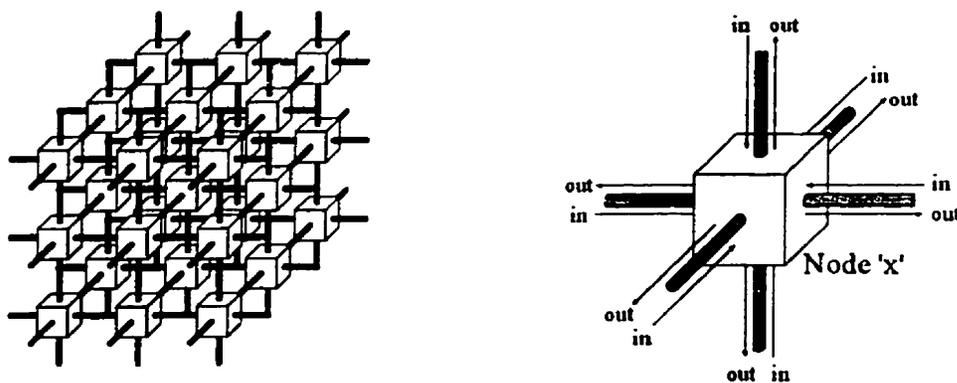


Figure 2-4: a) A 3-D Mesh interconnect, b) a single node

control in the network in order to guide messages to their proper destinations. The asynchronous transfer mode (ATM) protocol is one such control mechanism [ref 7].

The ATM protocol is based on a message passing technique. An ATM message is called a *packet* and has essentially two parts, a data payload and a destination address. A node in an ATM network moves packets by comparing the destination address of the packet with a set of addresses stored in a look-up table within the node. It then re-directs the packet to a node closer to its destination depending on this comparison. This type of local control is called packet-switching. A complete and fixed path from origin to destination is not required with packet-switched networks because each node redirects the packets in the appropriate directions. Along with the self-routing mechanism of each node, a global control mechanism is also used. The global control in a packet-switched network does not affect the propagation of packets directly, it is used to detect failures and congestion in the network and suggests alternate routes by updating the address look-up tables at the affected nodes. Another mechanism that is also used in ATM networks is packet buffering. When a node receives input from several other nodes each requesting access to the same output, the packets are queued in a linear first-in first-out memory buffer (FIFO). Buffering packets inevitably leads to latency in the system, but it is unavoidable due to the lack of point-to-point connections. By storing incoming packets until a path is available, contention is avoided but latency is increased. The probability of contention within a network has been one of the most intensely studied topics in the field of telecommunications. Mathematical treatises, based on queuing theory, have been used in many network analyses references [ref 8]. The ability to predict and compensate for a virtually random use of resources in a network is difficult and algorithms are continuously being revised [ref 9].

There are several examples of optical interconnect techniques that have been used to construct multistage interconnection networks (MINs). As indicated above, the crossbar embedding requires a large amount of fan-out in order to supply all outputs with all inputs. However, there are virtually no electronic systems which can provide the number of point to point connections, at a reasonable cost, required by a crossbar interconnect. The ability of optical fan-out, using holographic elements, provides a mechanism of replicating N-inputs M-times and directing this data to all outputs [ref 10]. There are several examples of systems that have used optical fan-out to produce Banyan networks. One such example is the System 5 demonstrator built by AT&T Bell

Laboratories (now Lucent Technologies) [ref 11]. This system optically interconnects five stages in a two-dimensional Banyan network using two by one optoelectronic switching nodes and Fourier computer-generated holograms (CGH). The inputs and outputs of the five-stage system are two-dimensional matrix fiber-bundles. Each of the five optoelectronic chips has several optical inputs and several optical outputs. The relay optics from one chip to the next in the system use CGHs to fan-out the signal from an output to many inputs on the next chip. This allows a Banyan-type interconnect to be implemented. Another example of a system that uses optics to implement an ATM-switch Banyan network is that of the NTT optical network systems laboratory [ref 12]. This system was constructed inside a standard printed-circuit board chassis and used polarization optics to guide data from one PCB to the next.

2.3) The Hyperplane architecture

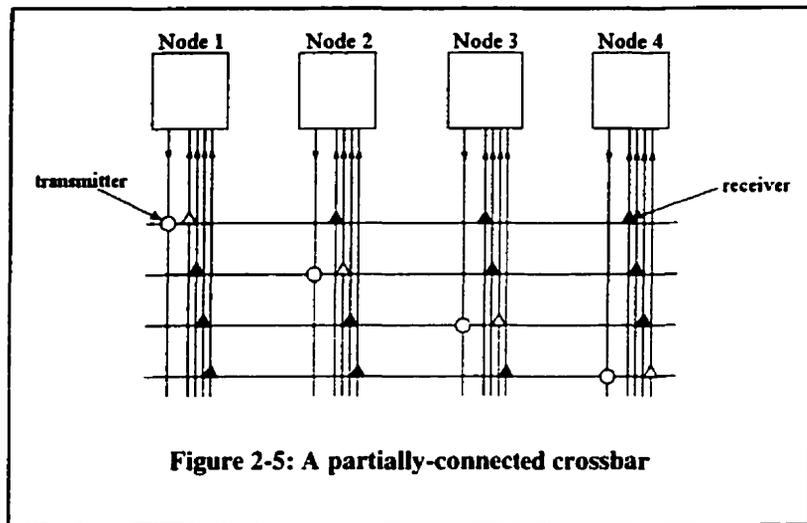
The generalized Hyperplane architecture can embed virtually any interconnect topology into the physical layer. The Mesh, the HyperCube, and the crossbar can be mapped into the interconnect for any intended application using the Hyperplane architecture [ref 4]. The reconfigurable nature of this architecture is advantageous because one embedding may be more efficient than another embedding for a particular application. For example, the 3-D Mesh interconnect may be well suited for local processing elements such as those found in massively parallel processors, but it may cause high latency in a application requiring broadcast capabilities. An analysis of the blocking probabilities for several different embeddings based on queuing theory have been done in other works [ref 13]. The generalized Hyperplane architecture also uses a combination of massive parallelism and time-division multiplexing (TDM). The parallelism is implemented using a large number of parallel optical data paths, and the temporal division is implemented using more intricate synchronization and a method analogous to instruction and data pipelining in microprocessors (see Chapter 6 - Synchronization) [ref 14].

The Hyperplane architecture presented in this chapter will be limited to a specific embedding: the sender-reserve partial crossbar. The partial crossbar differs from the full crossbar in terms of the number of directly accessible outputs. In a full crossbar, all

outputs are accessible at all times. Each node must have at least one input port and $N-1$ output ports assuming N nodes. For example, a fully connected crossbar with 32 nodes and 32 channels, each with a 32-bit data path, would require at *very least* 1024 electrical bond pads around the VLSI-OE chip at every node. The partial crossbar allows the number of outputs to be reduced to a reasonable number. Each of the N nodes in the system is provided with a single input port and a single output port. The partial crossbar will, at times, suffer from contention due to competition for the same output, but it is a far more realistic to implement, since only 64 bond pads would be required.

The optical implementation of the sender-reserve partial crossbar interconnect provides each node with one transmit channel. If there are N nodes in the system, there are N optical channels.

Therefore, each node has one optical transmit channel and $N-1$ optical receive channels [figure 2-5]. Because the interconnect is implemented optically, the large number of point to point



connections can still be accommodated. Once the optical data from each channel has been re-converted into electrical data within the node, one of the $N-1$ receive channels is selected and its data is routed to the electrical output port. Other channels that may have required access to the electrical output port are blocked from immediate access. The large number of optical channels directed at each node cannot all simultaneously gain access to the electrical output port. However, by selecting which optical channel gains access to the electrical output port, the total amount of hardware is reduced to a reasonable amount. This method is not an unreasonable tactic. An individual node should not be overburdened with output accesses at any one time. The probability that all nodes in the system will simultaneously transmit to a single node is very low. The loading on the network will normally be such that only a few channels will be attempting to gain access

to the output of a particular node at any one time. Some contention will occur, but this can easily be accommodated with an appropriate memory buffer.

Unlike the optical interconnect networks built by AT&T or NTT, the partial crossbar interconnect does not require computer generated holograms. Optical elements are used to provide many point to point connections and the data is not optically replicated. A CGH usually performs an optical fan-out, which means that an optically encoded message is replicated and re-directed to many spatially separated points. The CGH is ideal when trying to implement a Banyan-type interconnect with the butterfly connection pattern. The crossbar technique simply uses the two-dimensional surface area of the transceiver device to relay an entire image from one place to the next using optics [figure 2-6].

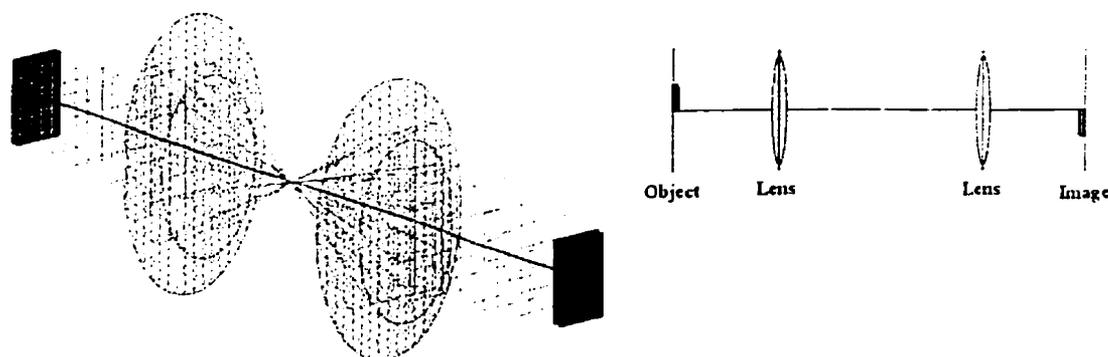


Figure 2-6: Imaging 2 planes using a pair of lenses

2.4) The smart pixel array

2.4.1) Introduction

Each Hyperplane node is an optoelectronic-processing chip called a smart pixel array. The smart pixel array is a hybrid microchip with optoelectronic devices integrated onto its top surface. Silicon CMOS is typically used as the processing part of the smart pixel array because complicated functionality can be implemented with relative ease. Gallium arsenide (GaAs) is typically used to form the optoelectronic devices because GaAs is a direct band gap semiconductor that can be structured into many different kinds of light emitting and light modulating devices as well as light detecting devices with high responsivities. The different types of optoelectronic devices and the processing required to implement these devices with silicon will be covered in a subsequent chapter.

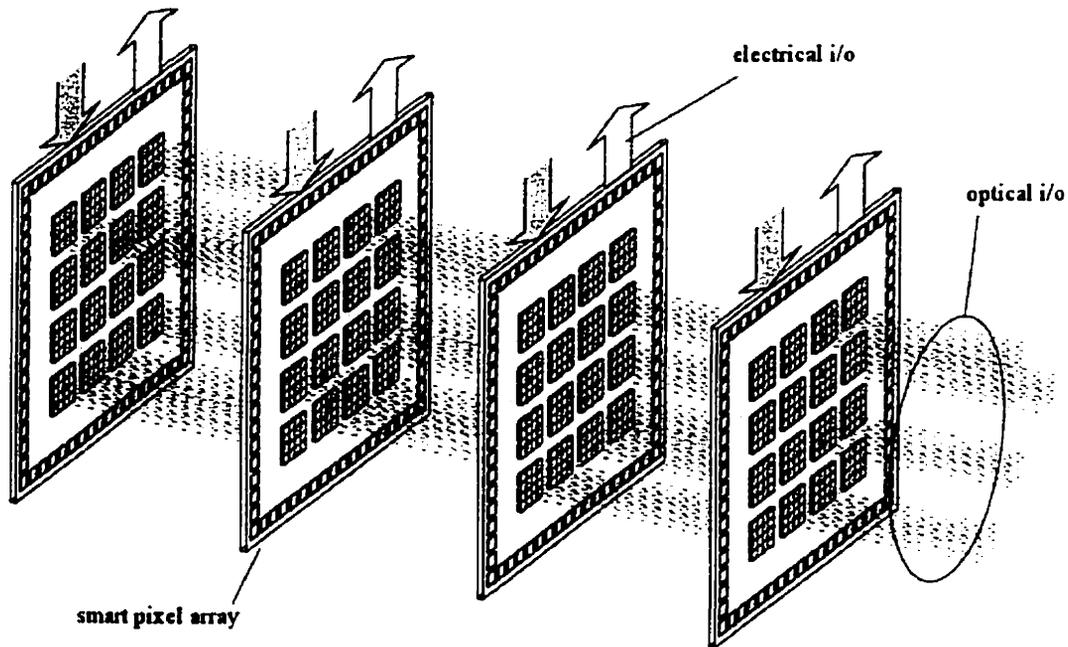


Figure 2-7: Concept picture of multiple SPAs with optical interconnects

The smart pixel array is the interface between the optics and the electronics [figure 2-7]. External electronics interact with the smart pixel array through a set of control lines and at least one electrical input port and one electrical output port. The

widths of the electrical input and output ports are made as large as possible to be compatible with the current trends in digital processing. The amount of electrical bandwidth to and from the smart pixel array may seem low compared to the optical bandwidth but is sufficient for most processing and switching architectures. Most high-speed bus structures in use today permit at most 64-bit word lengths; therefore similar word lengths from the smart pixel array are justified. There are on average 32 PCBs in a typical bus or backplane system. Therefore, if a crossbar interconnect is assumed, the whole system would require 32 optical channels of 64-bits on each smart pixel array. Thus, the smart pixel array may have over 2000 optical paths on the surface area of the chip, yet only 200 electrical paths from the perimeter of the chip. This is almost a 20 to 1 ratio of optical to electrical data paths. It is because of this large ratio that the smart pixel array requires switching electronics to maximize the use of the interface.

2.4.2) Channel definition

For a system with N nodes, the smart pixel array (SPA) must have at least N optical channels. The channels are parallel optical data paths M-bits wide (or M-smart pixels) and are equal in width to the electrical input and output ports around the perimeter of the chip. Each smart

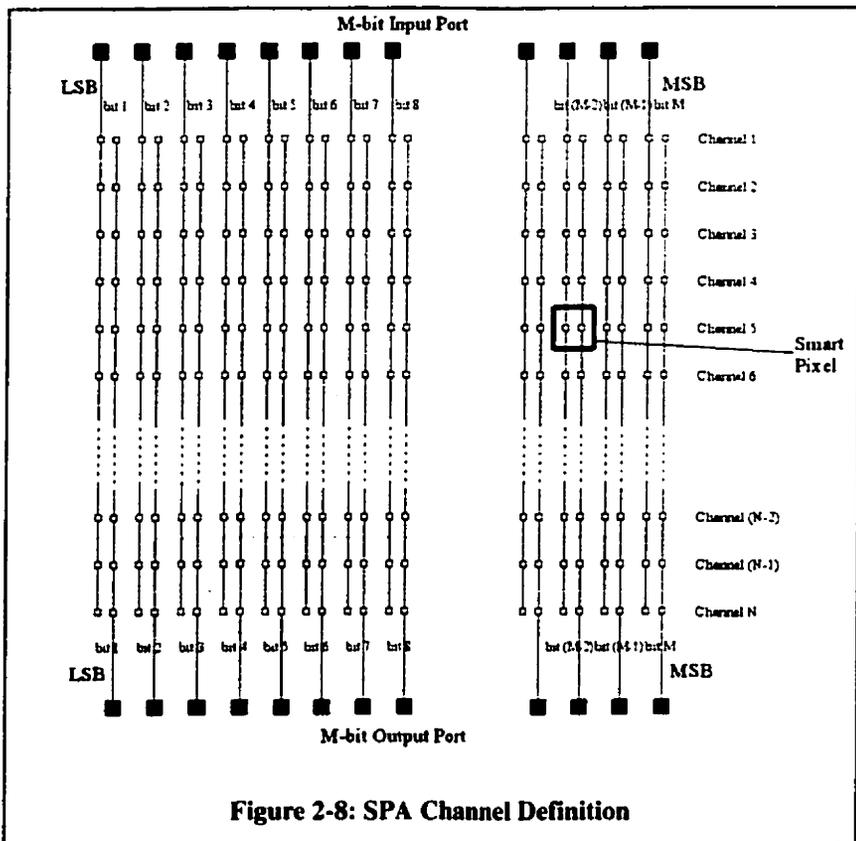


Figure 2-8: SPA Channel Definition

pixel array has one electrical input port of M-bits and one electrical output port of M-bits

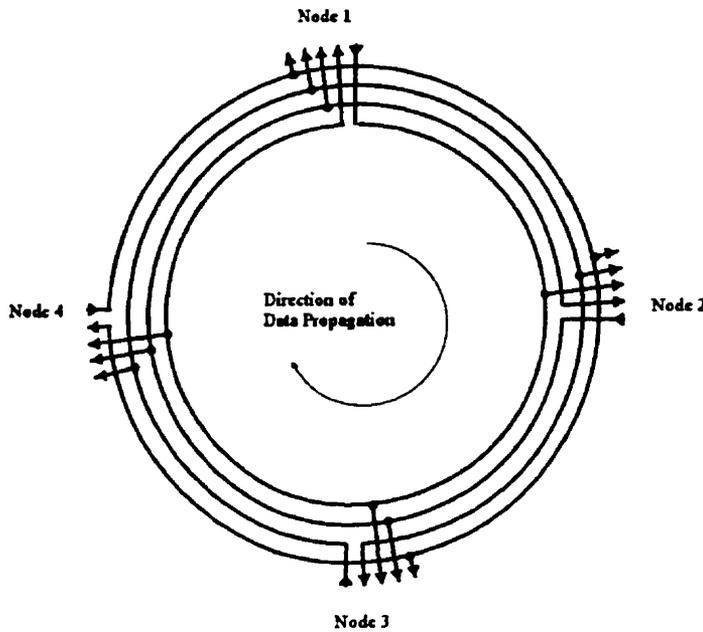


Figure 2-9: Concept of a uni-directional closed loop interconnect

which it uses to interface between the optical interconnect and the processing electronics. An electrical input and output port, each M-bits wide, are distributed to each optical channel via two M-bit electrical bus that passes vertically through each horizontal channel of the smart pixel array [figure 2-8].

2.4.3) SPA Protocol

The interconnect is based on an uni-directional closed ring, this is a consequence of the optical design and is discussed in Chapter 3. When data is introduced into the system, it travels from one node to the next in the same direction and arrives at the point it was sent. Even though data passes in only one direction, because the interconnect is a closed ring the data will circulate until it reaches its destination no matter where the data is introduced [figure 2-9].

The smart pixel array architecture combines some of the properties of circuit-switched networks with some of the properties of packet-switched networks. A smart pixel packet [figure 2-

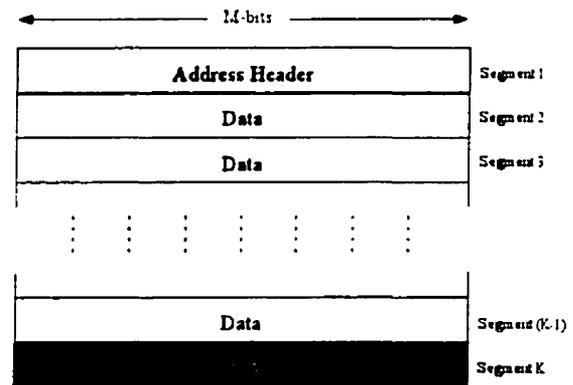


Figure 2-10: SPA packet

10], similar to an ATM packet [ref 15], is issued by one node and sent via an optical channel to all the other nodes in the system. The smart pixel packets are composed of K segments of M-bits and form a packet size of KM-bits. The first segment is composed of an M-bit address header, and is followed by K-1 segments of data. As the smart pixel

header travels the closed ring, it configures each smart pixel array in the system so that a fixed circuit exists between two smart pixel arrays for the duration of the packet transmission. The address header is used by each smart pixel array to determine if the packet should be captured. The header is compared to a permanent address assigned to each smart pixel array.

If two (or more) smart pixel arrays are attempting to simultaneously transmit data to a single smart pixel array, one of the two sender smart pixel arrays will not get access to the output port. Only one channel at a time can gain access to the electrical output port of a smart pixel array because of its limited electrical connectivity. The transmitted data will either have to be buffered at the receiver or re-transmitted by the sender. Obviously, an acknowledgement protocol is required for this type of transmission, but this is beyond the scope of this thesis and is covered in other works on telecommunication protocols [ref 8].

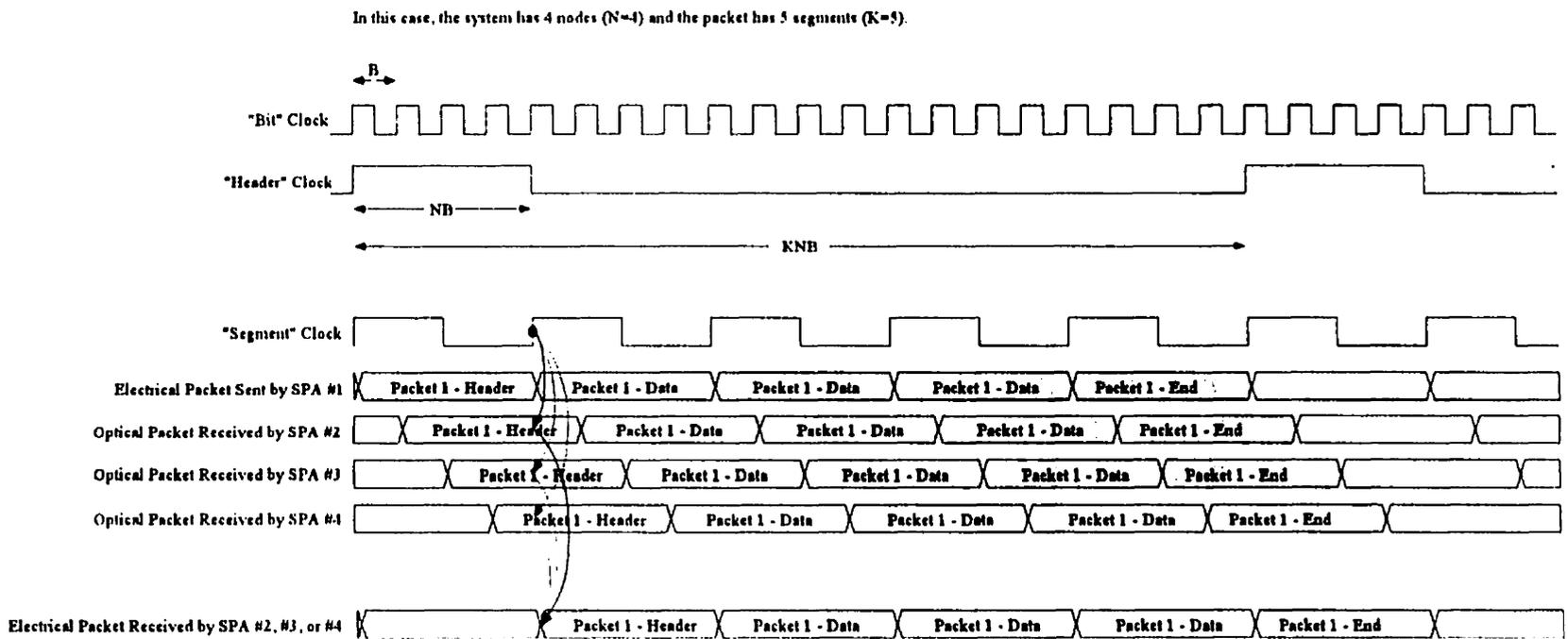
2.4.4) SPA clocking signals

One of the most challenging aspects of this architecture was the synchronization and control of the system as a whole. There were two methods that could be used to design a message-passing protocol for this system. A complex hand-shaking protocol could be used that would initiate communications, correct for lost data, and terminate the transmission - similar to that of an electronic bus. Although this protocol is reliable, it suffers from rather long latency. Another possibility was to use well-timed clock signals to sequence events. A set of well-timed clock signals can be used in place of a protocol if all the events take place in the correct order; similar to the method used in most modern microprocessors.

The figure below [figure 2-11] shows a typical set of waveforms used to pass the data around the interconnect. The primary clock has a period of B . The header-clock, which is responsible for indicating the "time-slot" of a packet header, has a period of KNB ; where K is the number of segments in the packet, and N is the number of nodes in the system. The header-clock is active during the header segment for a duration of NB seconds. Finally, the segment clock is used to continually sample the data at each node as it is passed from node to node.

These waveforms are only representative of the actual clocking scheme required by a system of this complexity. Issues such as skew and latency of both the data and the clocks must be considered (see Chapter 6 – Synchronization), this may influence the strategy for timing the nodes. Alternative and more complex timing strategies are described in the references [ref 4].

Figure 2-11: Waveforms for SPA packet routing



2.4.5) SPA logic

The smart pixel was the fundamental unit of the smart pixel array and was the transceiver for a single bit of data. For a system with N channels, where each channel was M-bits wide, there were MN smart pixels in the array. The smart pixel was responsible for both the optical-to-electrical and the electrical-to-optical conversion of data. It accessed four data paths; the electrical input data, the electrical output data, the optical input data and the optical output data [figure 2-12]. Only the electrical output data path had significant restrictions on the flow of data, and was the reason that the crossbar embedding was only a partial implementation.

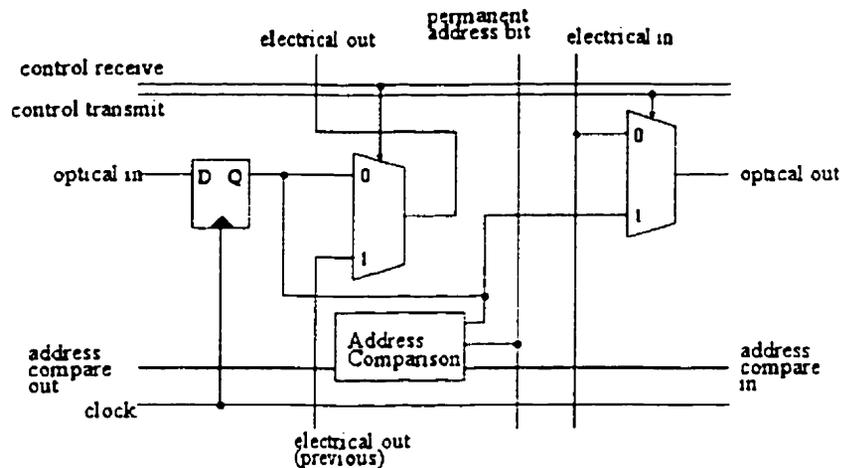


Figure 2-12: A smart pixel circuit diagram

The smart pixel had three primary modes of operation. The inject state, the extract state and the transparent state [figure 2-13]. When the inject state was active, the electrical input data to the chip was converted into the optical output data. When the extract state was active, the optical input data was converted into the electrical output data and routed to the electrical output port of the chip. The transparent state allowed optical input data to pass directly to the optical output with a minimum of electrical processing between.

There were six functional blocks in the smart pixel that carried out a few very simple functions at high data rates. The optoelectronic devices on the smart pixel were based on GaAs multiple-quantum well (MQW) P-i-N diodes. The structure of the optical transmitter depended on the nature of the optoelectronic device, but typically was as simple as a set of staged CMOS inverters. The optical receiver was also highly dependant on the type of optoelectronic detector used, but was typically based on a trans-impedance amplifier. The specific modulator and detector circuitry will be discussed in a subsequent chapter, and details of several circuits will be given (see Chapter 4 - VLSI Optoelectronics). The rest of the circuitry within the smart pixel was able to redirect data to the appropriate ports. To achieve this, two 2-to-1 multiplexers were required. The 'transmit' multiplexer was attached to the optical output and choose between the electrical input and the optical input. The 'receiver' multiplexer helped form the cascaded output data path through the channels to the electrical data output port. Finally, the smart pixel also required some form of synchronization to perform address recognition. Therefore, the smart pixel typically contained a delay-element, like a D-Flip Flop (a single bit memory element), that was able to store a bit during which the data could be compared to the permanent address in the chip.

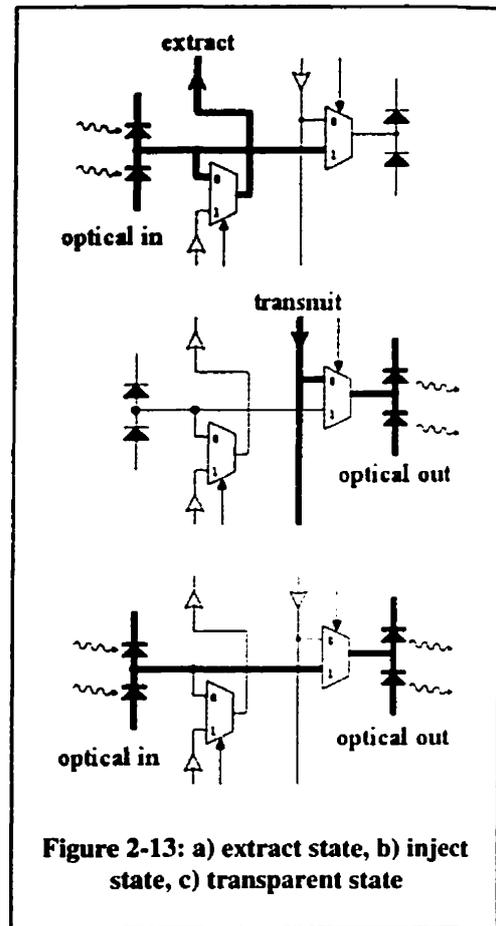
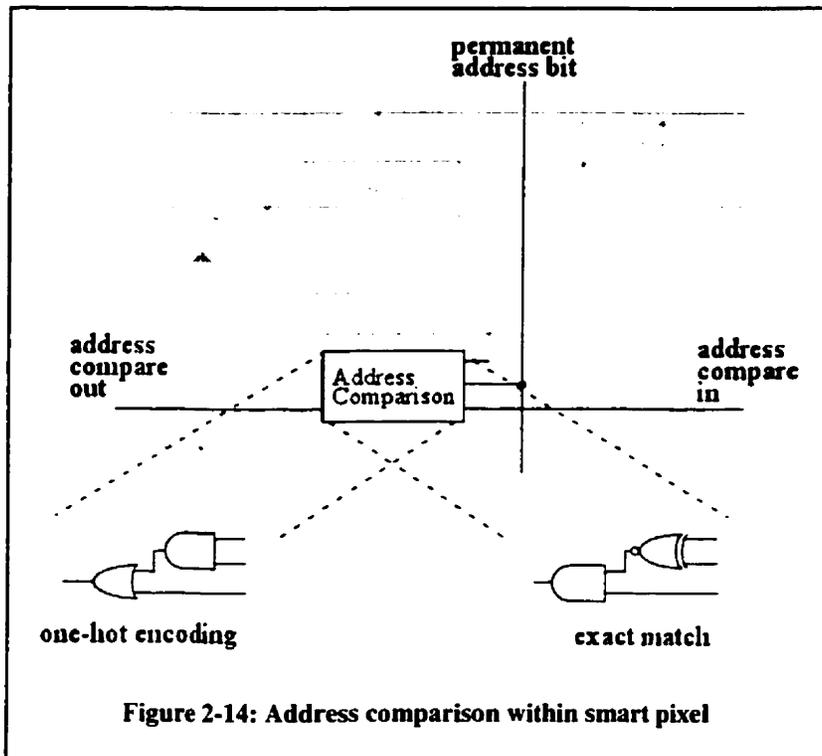


Figure 2-13: a) extract state, b) inject state, c) transparent state

The D flip-flop was placed between the optical input and the optical output within each smart pixel to regulate the flow of data and to hold the address header of a packet so that it could be compared with the permanent address of the smart pixel array. The permanent address was stored within each smart pixel array where each smart pixel was responsible for comparing one optical input bit with one bit in the permanent address. The results in each smart pixel were then cascaded through a chain of gates to produce



the address match signal. The smart pixel circuitry could contain circuitry that would detect either an exact match or one based on one-hot-encoding for selective broadcasting. The exact match required an exclusive-OR gate to compare the input optical data with the permanent address. Each exclusive-OR gate output would be

cascaded through a series of AND gates to produce the address match signal. The one-hot-encoding technique required an AND gate to compare optical input with the permanent address and a series of OR gates to produce the address match signal [figure 2-14]. The one-hot-encoding technique allowed for broadcasting as well as some flexibility in the implementation of the packet structure. The address match signal generated by the matching circuitry was used to configure the data paths within the smart pixel array.

2.4.6) Data-paths of the SPA

As described earlier, there was one electrical input port and one electrical output port, each M-bits wide. The electrical input port was connected to the “*transmit bus*” which was composed of M vertically routed signal lines that passed through each channel of the smart pixel array [figure 2-15]. The transmit multiplexer within each smart pixel could either use the data from the transmit bus or use the data from its own circuitry. The electrical output port was slightly more complicated because the N optical channels had only a single electrical output port to access. The receive control mechanism had to be

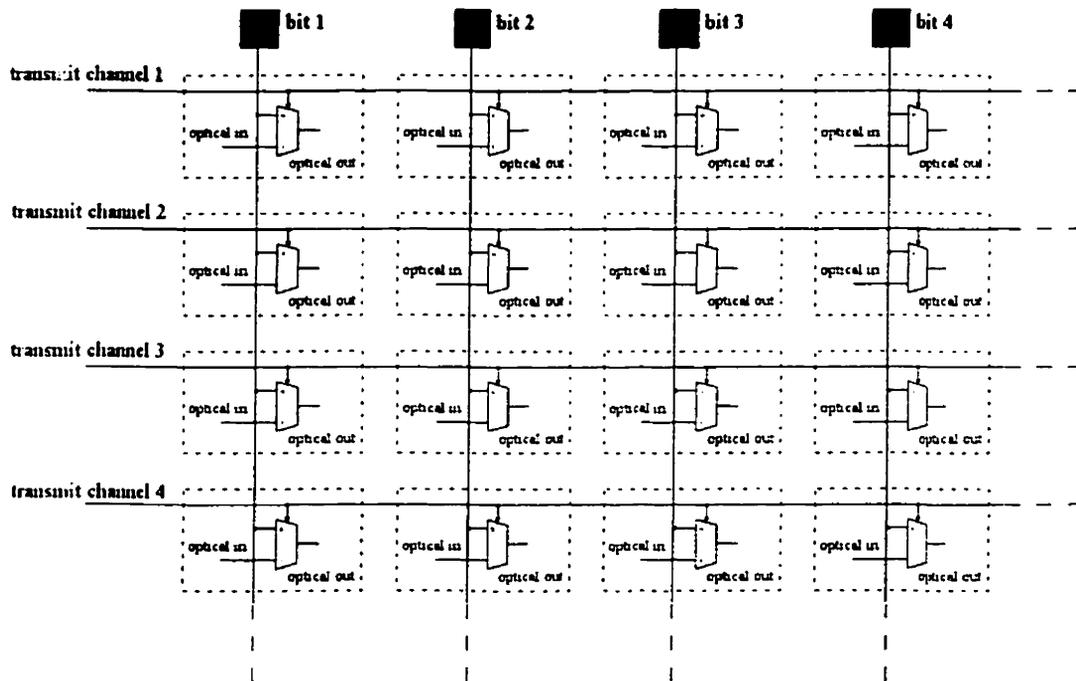


Figure 2-15: Transmit tree of SPA

(Each mux represents the transmit-mux of a smart pixel)

capable of handling any channel that requested the electrical output port of the chip with a minimum amount of contention. To access the electrical output port, each channel had to use a multiplexers connected to an output concentrator. This method allowed one electrical output port to serve N optical channels. The output concentrator was simply a parallel set of cascaded 2-to-1 multiplexers, called the receive-multiplexers, that passed

through each smart pixel in the array [figure 2-16]. When a channel was actively passing data to the electrical output port, the receive-multiplexer of that channel would direct data up through the cascade, the multiplexers in the above channels would be placed in a passive state and the data would eventually reach the electrical output port.

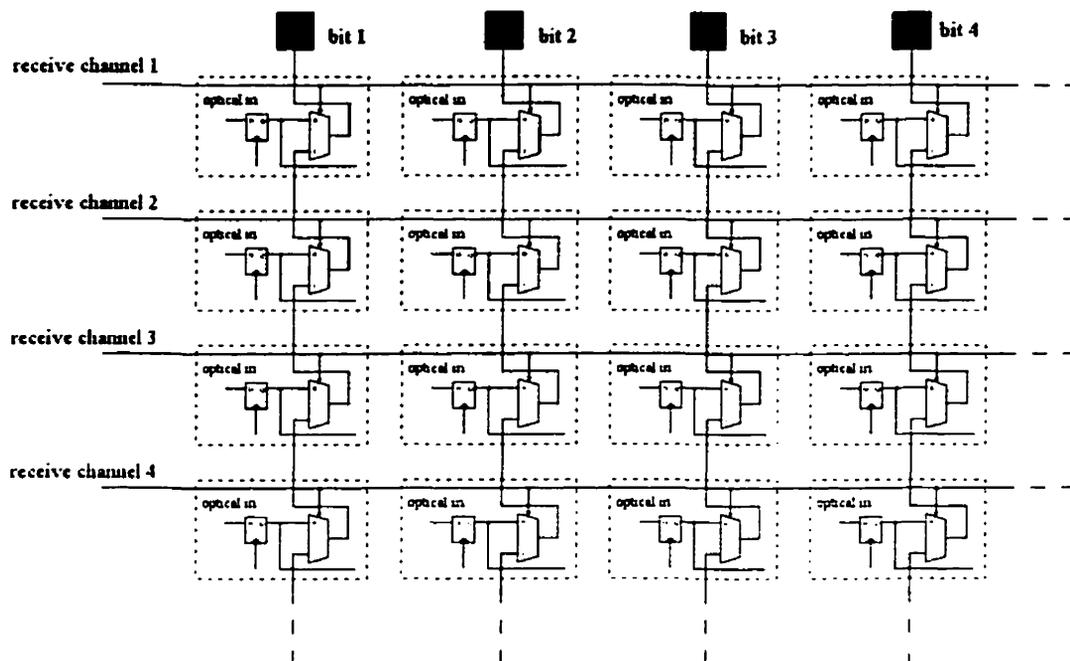


Figure 2-16: Output concentrator for a SPA

2.5) A simplified SPA

A simplified smart pixel array is shown below [figure 2-17]. The array has four channels where each channel contains four bits. All control bond pads, clock bond pads and data ports are shown including the register circuitry required to hold the permanent address and its associated bond pads. The address recognition circuitry included in this schematic is based on the one-hot encoding technique because it is one of the more general matching circuits. The schematic lacks any lines or bond pads that might be required for supplying power and ground and does not include the voltage biasing lines

for the optoelectronic devices. Also, the schematic includes neither the optical transceiver circuitry nor the optoelectronic devices, these structures are represented by darkened squares and labeled either "O-IN" for optical-input or "O-OUT" for optical-output.

The four-channel smart pixel array schematic is easily extendable to a $M \times N$ smart pixel array; the number of perimeter bond pads scales reasonably well with the size of the array. The array remains very generalized if control can be kept off-chip. However, a significant decrease in the number of bond pads can be obtained if the control circuitry which links the address recognition circuitry with the transmit and receive enable lines can be done on-chip. Incorporating the control circuitry within the chip may also increase the speed at which the channels are set-up because the signals would remain on-chip.

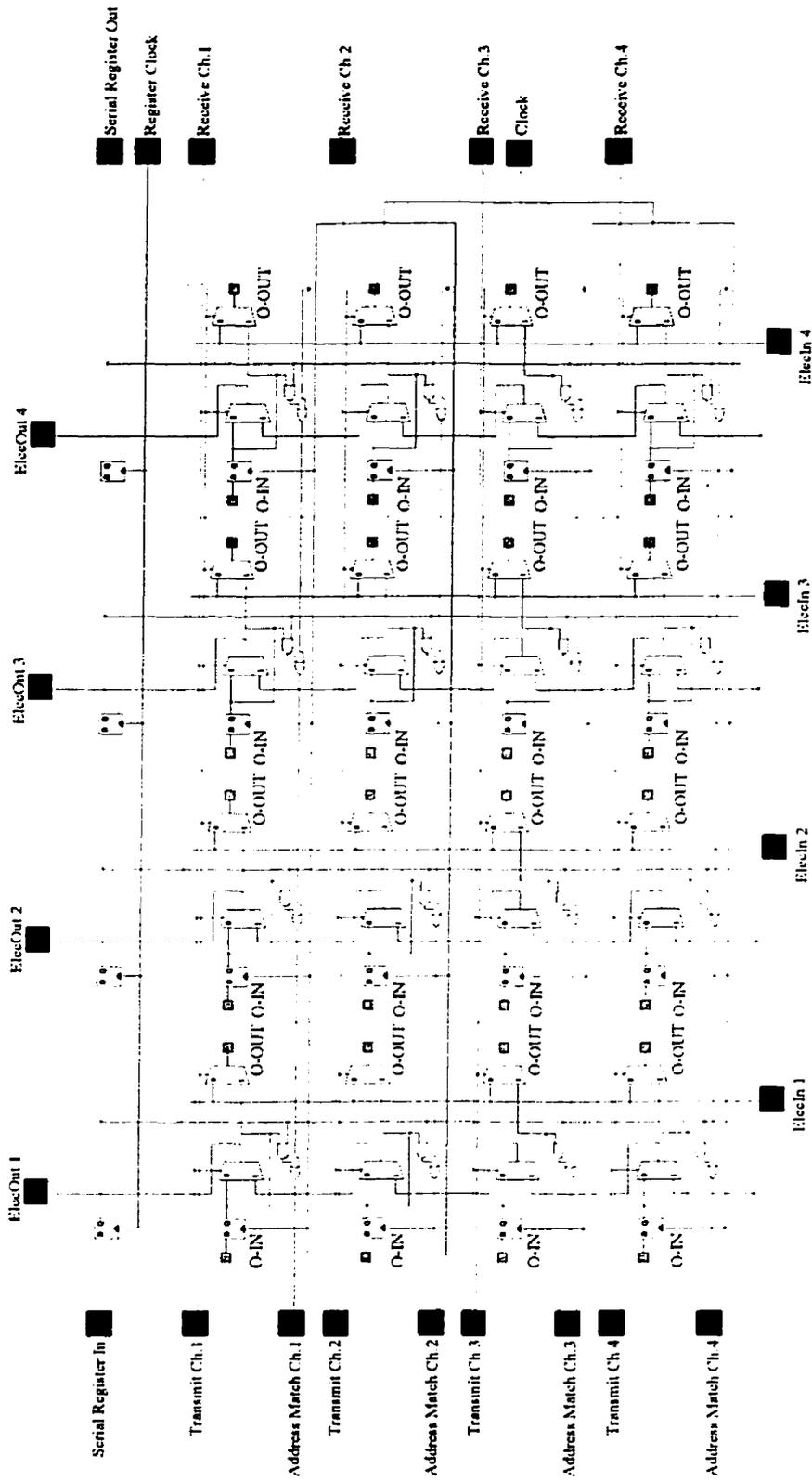


Figure 2-17: Digital schematic of 4x4 SPA

2.6) Conclusion

In this chapter, a method that was able to take advantage of both the three-dimensional propagation of guided light beams and the processing power of silicon CMOS microelectronics was proposed. A version of the crossbar interconnect was discussed as a particular embedding for the Hyperplane architecture, and the processing circuitry of the smart pixel array was shown.

In the following chapters, the hardware required to implement a complete optical backplane system will be discussed. These include the design of the optical interconnect, the design and testing of the optoelectronic microchip, and the method of synchronizing the entire system. The main focus of the following chapters deals with a particular implementation of the Hyperplane architecture for use in a free-space optical backplane demonstration system. However, the designs and results presented for each aspect of the system can be considered as “stand-alone” pieces of work that investigate particular aspects of the design of free-space interconnects.

2.7) References

- [1] T. Ono, Y. Yano, "Key technologies for terabit/second WDM systems with high spectral efficiency of over 1 bit/s/Hz", *IEEE Journal of Quantum Electronics*, Vol. 34, No. 11, Nov. 1998, pp. 2080-2088
- [2] F.M. Mousavi, K. Kikuchi, "Performance limit of long-distance WDM dispersion-managed transmission system using higher order dispersion compensation fibers", *IEEE Photonics Technology Letters*, Vol. 11, No. 5, May 1999, pp. 608-610
- [3] Park Jae-Hyun; Yoon Hyunsoo; Lee Heung-Kyu, "The deflection self-routing Banyan network: A large-scale ATM switch using the fully adaptive self-routing and its performance analyses", *IEEE/ACM Transactions on Networking*, Vol. 7, No. 4, Aug. 1999, pp. 588-604
- [4] T. Szymanski, H.S. Hinton, Optoelectronic smart pixel array for a reconfigurable intelligent optical backplane, United States Patent # 6,016,211, Issued Jan 18, 2000.
- [5] V.S. Adve, M.K. Vernon, "Performance analysis of mesh interconnect networks with deterministic routing", *IEEE Transactions on Parallel and Distributed Systems*, Vol. 5, No. 3, March 1994, pp. 225-246
- [6] J.P. Hayes, T. Mudge, "Hypercube supercomputers", *Proceedings of the IEEE*, Vol. 77, No. 12, Dec. 1989, pp. 1829-1841
- [7] M. Jeffrey, "Asynchronous transfer mode: the ultimate broadband solution", *Electronics & Communication Engineering Journal*, Vol. 6, No. 3, June 1994, pp. 143-151
- [8] D.P. Bertsekas, *Data networks*, Englewood Cliffs, N.J., Prentice Hall, 1992.
- [9] B.E. Ambrose, R.M. Goodman, "Neural networks applied to traffic management in telephone networks", *Proceedings of the IEEE*, Vol. 84, No. 10, Oct. 1996, pp. 1421-1429.
- [10] V.N. Morozov, W.T. Cathey, "Practical speed limits of free-space global holographic interconnects: time skew, jitter and turn-on delay", *Applied Optics*, Vol. 33, No. 8, March 1994, pp. 1380-1390.
- [11] F.B. McCormick, T.J. Cloonan, A.L. Lentine, J.M. Sasian, R.L. Morrison, M.G. Beckman, S.L. Walker, M.J. Wojcik, S.J. Hinterlong, R.J. Crisci, R.A. Novotny, H.S. Hinton, "Five-stage free-space optical switching network with field-effect transistor self-electro-optic-effect devices", *Applied Optics*, Vol. 33, 1993, pp. 5153-5171.
- [12] K. Hirabayashi, T. Yamamoto, S. Hino, Y. Kohama, K. Tateno, "Optical beam direction compensating system for board-to-board free space optical interconnection in high-capacity ATM switch", *Journal of Lightwave Technology*, Vol. 15, No. 5, May 1997, pp. 874-882
- [13] Manoj Verghese, "A software based design space exploration of a free-space photonic backplane" McGill University, Montreal, Canada, 1995.
- [14] K.A. Sakallah, T.N. Mudge, T.M. Burks, E.S. Davidson, "Synchronization of pipelines", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 12, No. 8, Aug. 1993, pp. 1132-1146
- [15] X. Gu, K. Sohraby, D.R. Vaman, *Control and performance in packet, circuit, and ATM networks*, Kluwer Academic Publishers, Boston, 1995

Chapter 3: Optical Interconnects

3.1) Introduction

There are two general fields of optical design that can be used to relay information from one place to another. The guided-wave approach, which uses elements such as optical fibers to force light along enclosed regions of space, and the free-space approach, which is based on individual optical elements such as lenses to redirect the light. The field of fiber optics is a very mature technology, especially in high-speed serial data communications and is discussed in numerous references [ref 1,2]. Conversely, most free-space optical design techniques have been used almost exclusively for imaging objects for human observers such as cameras or telescopes [ref 3]. However, in the past few years, research into free-space photonic interconnects has shown that free-space optics may hold the promise of high-speed massively parallel digital networks [ref 4].

In many respects the guided-wave approach is far easier to implement than the free-space approach because of the simplified optomechanical structures needed to hold fiber-based systems. A guided-wave system usually requires only the two ends of the fiber to be well aligned [figure 3-1], while the free-space method requires multiple elements to be aligned together. The Infineon Technologies PAROLI optical link [ref 5]

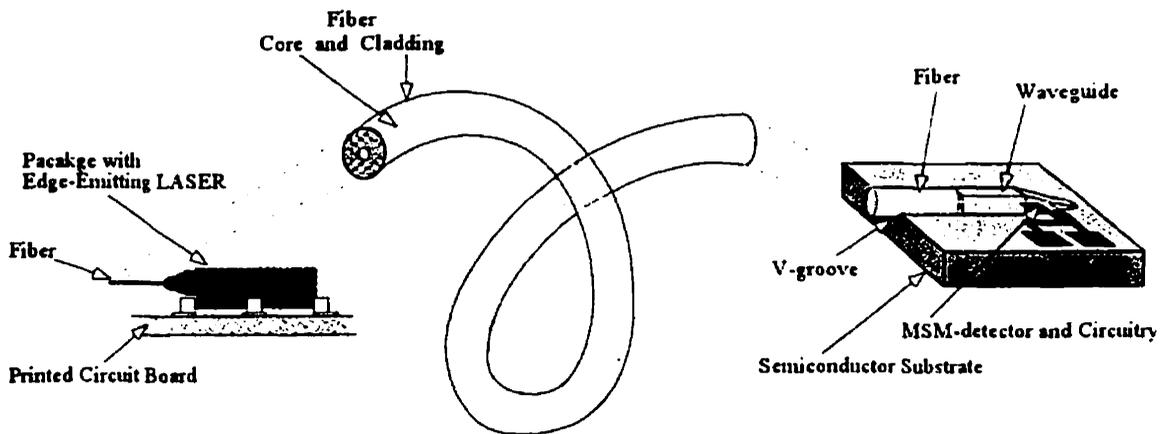


Figure 3-1: Typical single-fiber link

is an example of a transmitter-receiver system that has 12 2.5-Gbps parallel optical fibers linking a linear array of VCSELs to a linear array of P-i-N diodes using a graded-index multimode fiber ribbon [figure 3-2].

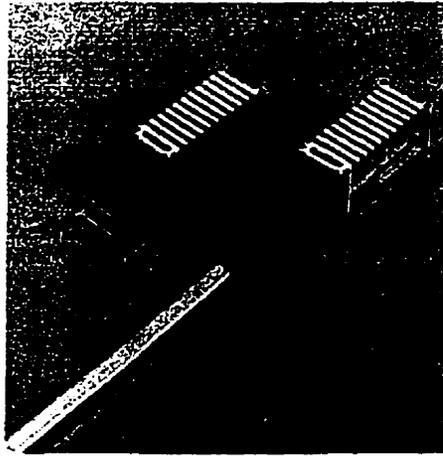


Figure 3-2: Infineon PAROLI fiber link

However, one of the drawbacks to fiber based systems, is that it may be difficult to assemble larger arrays of fibers within the compact volume of a typical printed circuit board chassis. This task is made especially difficult when the printed circuit boards are spaced only centimeters from each other. The guided-wave approach also does not allow for convenient fan-in or fan-out of data, and is predominantly used for point-to-point

interconnects.

As an alternative to the guided-wave technique, the free-space optical interconnect may be used. A very well known example of a simple free-space optical interconnect is the compact-disk (CD) optical pick-up head [figure 3-3] [ref 6].

The CD pick-up is a relatively simple design that tracks the reflection of a single modulated beam from the surface of a CD and relays the modulated signal onto a photodetector. Free-space optical interconnects offer many advantages. The free-space interconnect can perform many unique

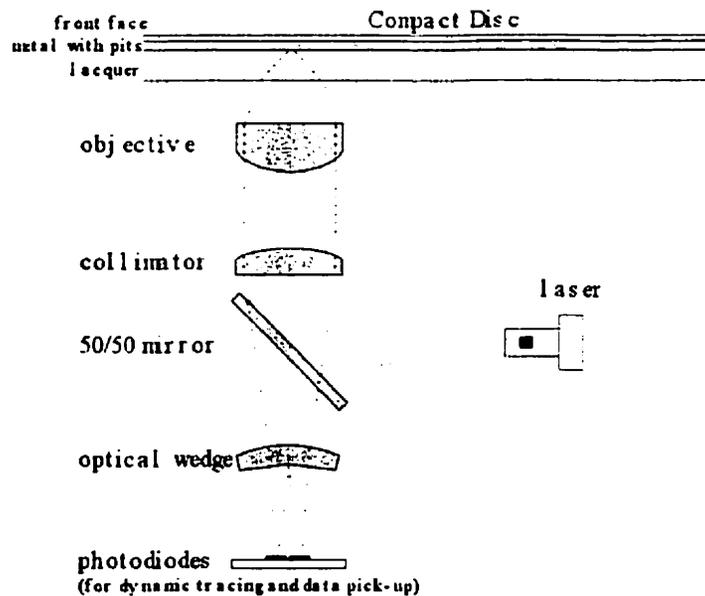


Figure 3-3: Typical optical CD pick-up

functions very easily, such as beam combination, data fan-in and fan-out, and it can provide an extremely high number of interconnects in a very small volume of space. However, a free-space interconnect also suffers from a number of design challenges. The ability to have a high tolerance to physical misalignment in conjunction with a cheap and easy method of holding each element (called optomechanics) is the greatest challenge. The number of components in a free-space design along with their reliability under adverse environmental conditions are also critical issues that arise.

This chapter will outline two free-space optical designs based on demonstration systems by the McGill Photonic Systems Group. These designs allow optical communication between multiple printed circuit boards, where the separation distance between PCBs is of the order of a few centimeters. It will begin by discussing some of the characteristics important to optical design by briefly explaining some of the basic principles behind certain optical components. This chapter is included primarily to describe the nature of two specific free-space optical interconnects, and show how the optoelectronic microchips use the free-space optical paths. A comprehensive and general overview of free-space optical interconnects will *not* be given here, nor will a detailed characterization of the McGill demonstration systems be given. Both of these topics span far too much information to be directly relevant to this thesis, and can be found in other works [ref 7, 8, 9].

3.2) Free-Space Optical Interconnects

3.2.1) Introduction

The objective behind creating a free-space optical interconnect was to provide a new type of telecommunication switching fabric that combined optics, microelectronics and optoelectronics. The switching fabric would implement the architecture outlined in a previous chapter (see Chapter 2 – Architecture), and would provide over 1000 optical connections per processing node. The typical number of nodes, and the overall functionality would be similar to any present-day fully electronic telecommunications switch, but provide roughly 1000 times the bandwidth. For example, the Centillion 1000 Series multiservice ATM switch from Bay Networks can handle multiple OC-12

connections with data rates of up to 10-Gbps [figure 3-4] [ref 10]. The free-space optical interconnect is intended to supply at least 1-Tbps switching capacity.

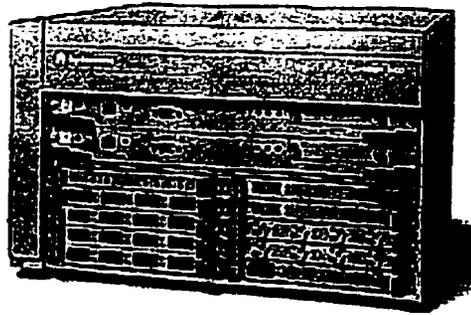


Figure 3-4: The Centillion ATM switch

An initial set of conditions was placed on the design of the optical interconnect so that a solid set of parameters could be outlined with which to build on. It was also important to target areas of the design that could enhance a telecommunications switch topology without starting from scratch.

To enhance the performance of a telecommunications switch, an enormous number of point-to-point (node-to-node) connections were required. All the nodes would also be required to communicate simultaneously (or at least a large percentage of them). This huge bandwidth requirement was needed, otherwise, there would be no benefit in exploring the optical alternative compared to the advances that are continuously made in electronic switching. To satisfy the general form-factor of today's systems, the size of the optical design had to be comparable with present-day electronic telecommunication switches. The optical interconnect was to replace the electronic bus of the system and was responsible for interconnecting multiple printed circuit boards (PCBs). The separation between PCBs in most electronic chassis is roughly 3-cm. Therefore, it was decided that this dimension would also be the target for the optical design. Another important goal was to have an optical system that was easily expandable. The optical design had to allow for any number of nodes in the system, at least up to the maximum number within an electronic chassis.

3.2.2) Interconnect topology

The interconnect topology had to guarantee that any node in the system could communicate with any other node in the system, this could be accomplished in one of two ways. The interconnect could be built in a linear fashion, with optical signals moving

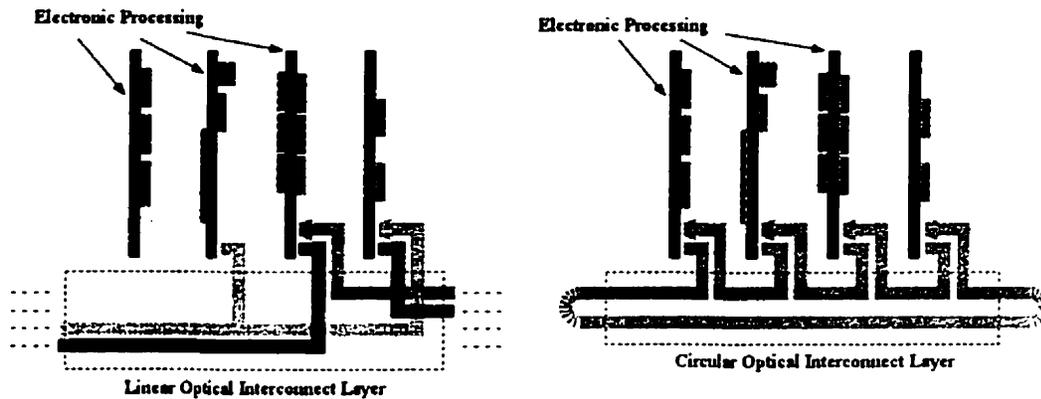


Figure 3-5: a) Concept of a linear interconnect, b) concept of a circular interconnect

up and down the optical path much like a typical electrical bus. Alternatively, the interconnect could be built in a ring configuration, with optical signals moving in only one direction but eventually reaching their destination as they traversed the ring [figure 3-5]. Although free-space optical designs can implement beam-combination and allow for fan-out and fan-in, the task of creating a bi-directional link is very difficult. The lack of bi-directionality in optical systems follows from the principle of reversibility [ref 11]. The principle of reversibility means that the path of light traced through a system in one direction will be identical to the path traced through the system when the light travels in the opposite direction [figure 3-6]. Therefore, a path that combines both in-coming and out-going beams that propagate in opposite directions must be able to overcome this property in some way. It is difficult to realize an optical element that is capable of directing light originating from the left and directing it downward as well as directing light originating from the right and directing it downward. This is made even more difficult assuming the node must be able to selectively send its data either to the left or the right. There have been many systems proposed that have been based on linear, bi-

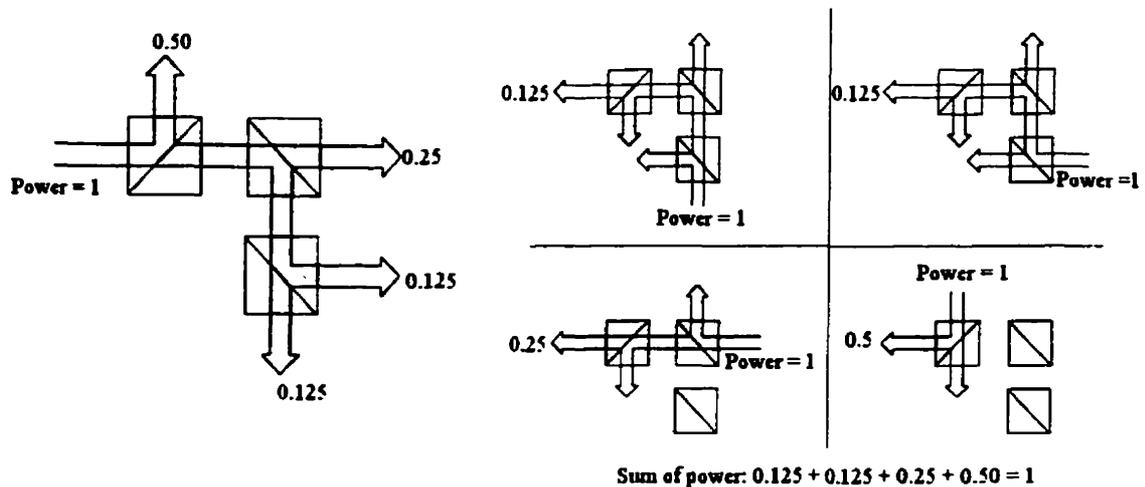


Figure 3-6: Concept of reversability

directional optical designs and these have used elements such as computer-generated holograms (CGHs), off-axis lenses, and diffraction gratings [ref 12,13].

An example of this type of bi-directional design is the one shown here [figure 3-7], it uses diffraction patterns on a glass substrate. This design attempts to interface multiple nodes with each other in a broadcast configuration by using selective diffraction patterns etched on to the top of the substrate. Unfortunately, this method can suffer from power non-uniformity as well as tolerance and alignment issues, not to mention a large cost.

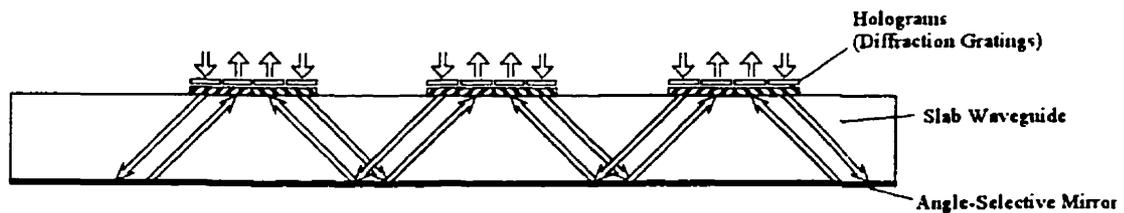


Figure 3-7: Diffractive slab-waveguide

As an alternative to the linear, bi-directional types of optical interconnects, a circular, uni-directional optical interconnect can be built. The uni-directional closed-ring is a simple alternative that maintains full connectivity among all nodes. The uni-directional closed-ring propagates all the beams in the same direction allowing all nodes to interrogate the data while greatly simplifying the type of optical components required.

Although there were several advantages and drawbacks with this type of optical relay with respect to the electronic architecture, one of the major drawbacks with the

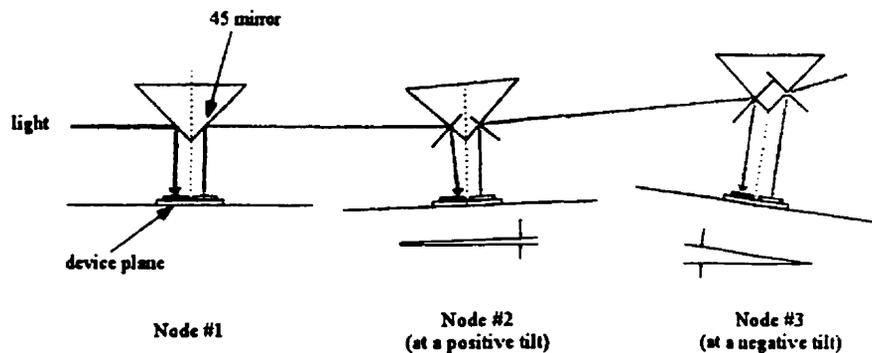


Figure 3-8: Cascaded linear misalignment with corrections

closed-ring approach was the possibility that the misalignments in the construction of the optical system would not allow the ring to be closed. Unlike a linear system, where small misalignment errors between nodes could be handled by subsequent stages in the system [figure 3-8], the closed-ring required virtually perfect accuracy. Since the initial misalignment and the corrections for these misalignments would cascade through the optical system. When the beams from the last stage were directed back to the first stage, the beams could be shifted and unfocused [figure 3-9].

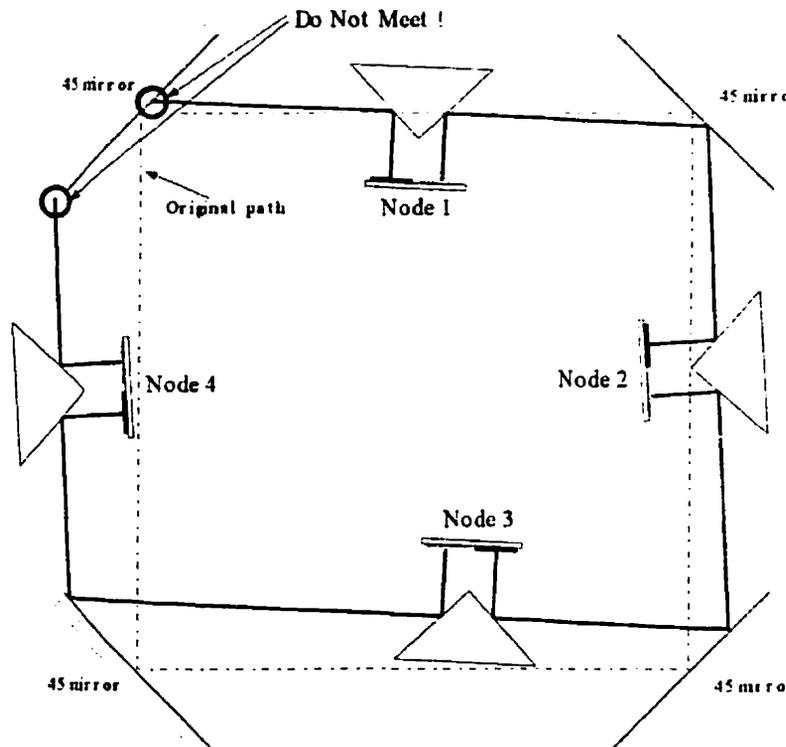


Figure 3-9: Misalignment of a circular optical system

3.2.3) Basic optical relay design

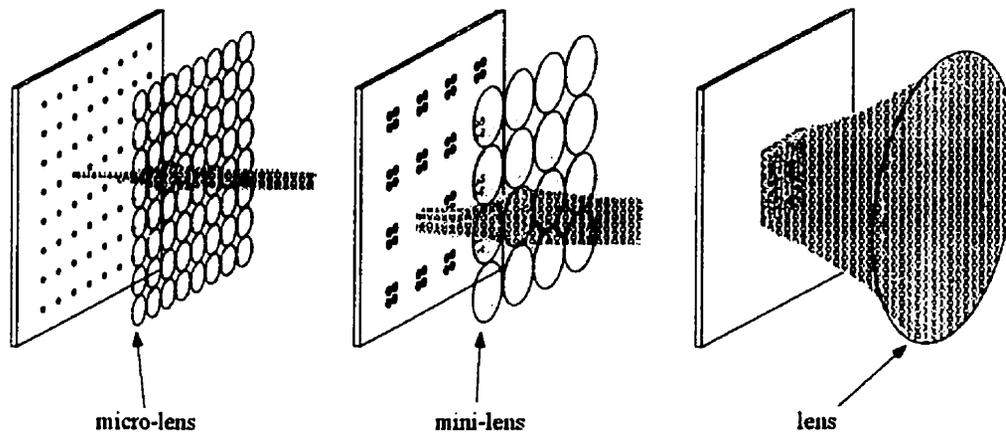


Figure 3-10: Three types of optical relay

There were several methods that were available to create an optical link capable of supporting thousands of optical beams. An enormous number of very small lenses arranged in a tightly packed array could be used, called a micro-lens array. A modest number of modest size lenses arranged in an array could be used, called a mini-lens array. And finally, a single, large, field lens could be used [figure 3-10]. The micro-lens array allowed only a single optical relay per lens. Therefore, the optoelectronics would have to be organized in a regular array to suit the micro-lens array. This type of optical design usually required a particular relay technique called the maximum-lens-to-waist configuration to provide sufficient working distances between micro-lens arrays and achieve a high density of signals. The mini-lens array was of the same form as the micro-lens array, except that each lens in the mini-lens array was larger and could support several optoelectronic devices. This technique is called clustering and is discussed later. The mini-lens array could support a relatively large working distance as well. The final relay technique available was to use very large lenses to image the entire field of view of the microchip. This design was the least attractive, because the lens had to have a very large flat field of view with very little aberrations over the given area. The relay distances also were slightly too large for any reasonable size lens. Therefore, this technique was reserved for special cases, such as hybrid optical links.

3.2.4) Optical modeling

The technique of imaging an object from one plane to another can be modeled with simple ray-tracing techniques [ref 14]. Ray tracing is based on the assumption that light emanating from an object can be represented as planar electromagnetic wavefronts. The rays are lines perpendicular to the wavefronts and indicate the direction of propagation. In the limit, spherical wavefronts can also be modeled as planar wavefronts and thus ray-tracing is again possible. Ray tracing and Snell's Law have led to equations such as the paraxial thin-lens equation that predict where the image of an object will exist after passing through a lens [Eqn. 3-1]. For example, an object placed at a distance s_1 from a lens will form an image at a distance s_2 from the other side of the lens related by the focal distance of the lens.

Eqn. 3-1: Thin Lens Equation

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f}$$

Unfortunately, a major drawback of light emitted by a point-source is that only a small percentage of optical power from the object can be captured by the lens and redirected toward the image plane. The light from a point source emanates in spherical wavefronts and cannot be confined to a specific region of space. The percentage of power from a point source that can be captured by a lens with diameter D and focal length f is given in equation [Eqn. 3-2]. However, another treatment of the propagation of light exists that is spatially localized, and is modeled using Gaussian beam propagation theory.

Eqn. 3-2: Percent Power

$$\% = \frac{100D}{2\pi\sqrt{D^2/4 + f^2}} \sin^{-1}\left(\frac{D}{2\sqrt{D^2/4 + f^2}}\right)$$

Gaussian beam propagation [ref 15, 16] was originally developed to model the behavior of laser cavities. The profile of the gain structure and the type of optical resonator in a laser gives rise to specific spatial modes of the output light. The lowest order mode is the TEM_{00x} mode and its radial intensity profile looks like a two-dimensional Gaussian curve [figure 3-11]. The homogenous partial differential equation for electromagnetic wave propagation remains the same [Eqn. 3-3]. However, unlike the sinusoidal solutions of a plane waves propagating in free-space [Eqn. 3-4], Gaussian

analysis assumes that the strength of the electric field tends to zero at large distances radially away from the axis of propagation. This assumption produces solutions that indicate that the beam is spatially confined to a region of space as it propagates [Eqn. 3-5]. The solutions also predict that the shape of the wavefront is not quite spherical nor is it planar. Strictly obeying theory, ray-

tracing techniques can no longer apply to Gaussian beams. However, ray tracing was used to model the approximate position of the Gaussian spots assuming that the Gaussian wavefronts were approximately planar near the paraxial axis. A particularly useful equation that comes out of the derivation can be used to calculate the beam-waist ω_0

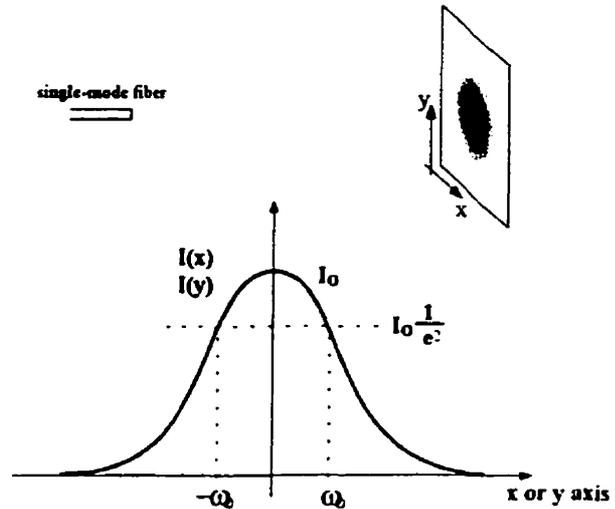


Figure 3-11: Gaussian beam profile (TEM 00)

Eqn. 3-3:
$$\nabla^2 E + \frac{4\pi^2 n^2}{\lambda^2} E = 0$$

Eqn. 3-4:
$$E = E_0 \exp(-jkz) = 0$$

Eqn. 3-5:
$$E = E_0 \exp(-jkz) \exp(-jP(z)) \exp\left(-\frac{kr^2}{2q(z)}\right)$$

where $E \rightarrow 0$ as $x, y \rightarrow \infty$

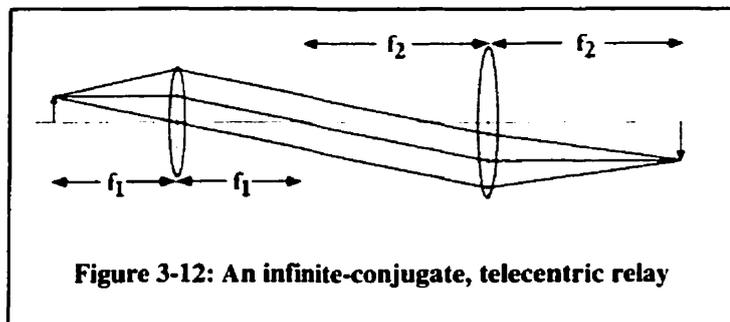
Eqn. 3-6:
$$\frac{1}{q(z)} = \frac{1}{R(z)} - j \frac{\lambda_0}{\pi \omega_L^2(z)}$$

and
$$\omega_L^2(z) = \omega_0^2 \left[1 + \left(\frac{\lambda_0 z}{\pi \omega_0^2} \right)^2 \right]$$

(also called the spot size) as a function of propagation distance as the beam expands due to its natural divergence [Eqn. 3-6]. The beam-waist indicates the normalized $1/e^2$ intensity points of the Gaussian curve, approximately 66% of the power is contained within this region. In most of the calculations, a region 3-times the beam-waist, $3\omega_0$, was used as the initial spot-size so that approximately 99% of the power could be contained within this region. The definition of minimum beam-waist (or spot-size) is the region in space where the field curvature of the Gaussian beam is infinite (i.e.: the spot-size is where the electromagnetic field is, momentarily, a plane wave). Therefore, using laser light and Gaussian beam analysis an optical system can be designed so that it is remarkably efficient and relays virtually all the optical power provided by the source.

There are two special cases of Gaussian beam propagation. The first is called the *maximum lens-to-waist configuration* and involves an optical system that has the minimum spot size a Rayleigh-range distance further from the focal plane of the lens. The second is called the *maximum beam-waist configuration* and involves a telecentric, infinite-conjugate system where the spots are at the focal plane of the lenses.

A telecentric, infinite-conjugate system involves a pair of lenses, with focal lengths f_1 and f_2 , that have been placed a distance $f_1 + f_2$ apart. This type of relay requires that all focal planes be coincident with one another [figure 3-12]. The non-telecentric optical relay is essentially any other configuration of object, lens and image. When a telecentric, infinite conjugate relay is used, the Gaussian model predicts that the wavefronts of the optical fields are planar at the focal planes of the lenses. To calculate the



focal length of a lens assuming Gaussian beam propagation [Eqn. 3-6] must be used. For example: if a lens with a diameter of $3\omega_L = 800\text{-}\mu\text{m}$ and a spot-size of $3\omega_0 = 25\text{-}\mu\text{m}$ is at the device window, the focal length f of the lens can be calculated. In this example, it was assumed that a 3ω diameter captures 99% of the power in a spot and that the lens had an aperture of 3ω , the wavelength was also assumed to be 850-nm. The focal length can be calculated to be $z = f = 8.21\text{-mm}$. When dealing with the maximum lens-to-waist

configuration, the same procedure is used, except that the initial spot is placed at one Rayleigh-range $z_r = \pi w_0^2/\lambda$ from the focal plane of the lens [figure 3-13]. For example: if a lens with a diameter of $3w_L = 125\text{-}\mu\text{m}$ and a spot-size of $3w_0 = 6.5\text{-}\mu\text{m}$ at the device window, the focal length f of the lens can be calculated. In this example, it is also assumed that the lens captures 99% of the power, and that the wavelength is 850-nm. Therefore, the focal length would be 315.86- μm with a Rayleigh-range length of 17.35- μm . To calculate the maximum lens-to-waist distance on the right side of the lens, the equation $f(1 + f/(2z_r))$ is used and results in a distance of 3.19-mm from the lens.

Although an optical system based on a maximum lens-to-waist configuration allows a greater spot density than a system based on the maximum beam-waist configuration, it can be much more difficult to align. The maximum beam waist configuration, or infinite-conjugate, telecentric system, can be more than 3 to 4 times more tolerant to angular misalignment for the types of optical systems considered in this thesis. However, the spot density for the infinite-conjugate, telecentric systems is roughly 3 to 4 times less. Although, with the device window clustering techniques discussed in section 3.3.3, this density can be significantly increased.

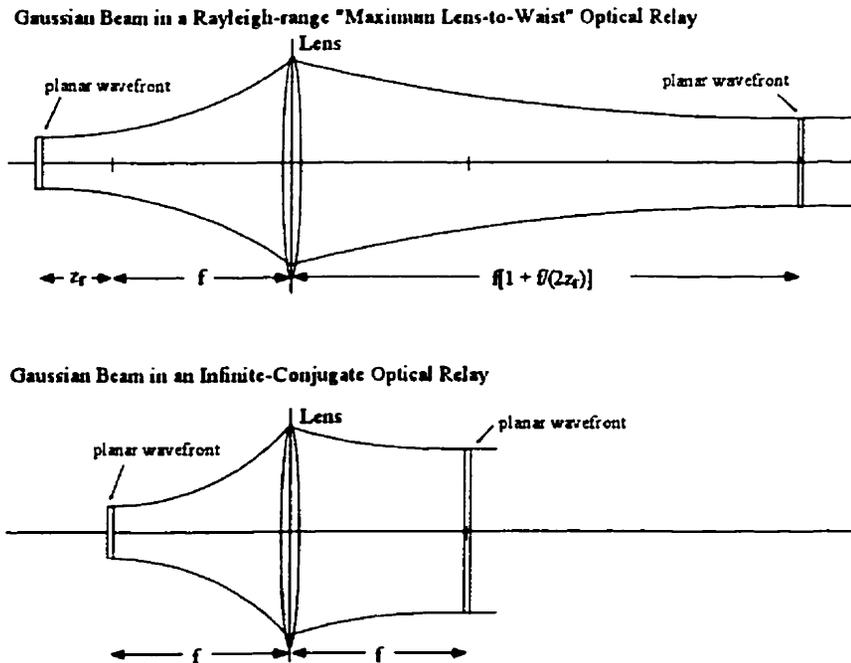


Figure 3-13: Gaussian beam in two types of relay

In the design of optical systems, both ray tracing and Gaussian beam propagation techniques are essential tools. However, factors such as the optical loss due to unwanted interface reflections and the loss due to non-ideal polarization states may impact the design tremendously. Finally, the most important characteristic of an optical design is its tolerance to mechanical misalignment. Models for the amount of power relayed by a link that has been misaligned laterally, longitudinally, or by some tilt are very important because they show areas of the design that are very sensitive. Tolerance analysis is done in several ways. The Gaussian models allow integration of areas to estimate the total power captured by a lens and they usually form solutions involving the error-function. Other methods, such as Fresnel- Kirchhoff scalar diffraction theory, allow optical aberrations and tolerances to be modeled more accurately [ref 17].

3.2.5) Beam combination optics

There were two types of optical components used in the design of the systems described in section III that allowed some form of directional control to be implemented. These elements were the polarized beam-splitter (PBS) and the quarter-wave plate (QWP).

When both of these elements are used together, along with a mirror at normal incidence, a beam combination technique is obtained. If the incoming light is horizontally polarized and directed at the PBS, the PBS will

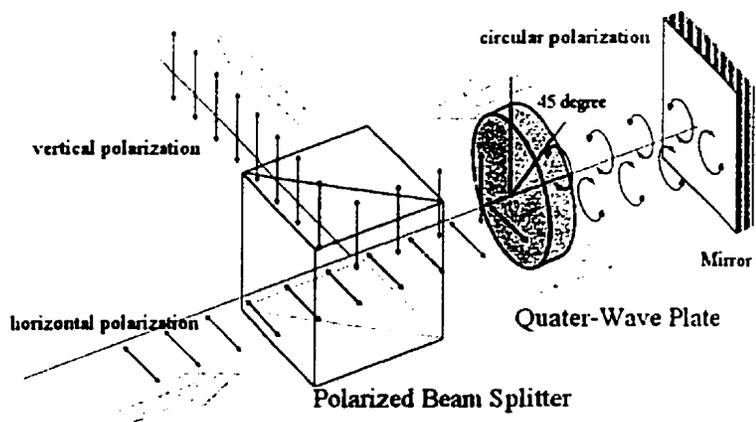


Figure 3-14: PBS+QWP assembly

allow the horizontally polarized light to pass straight through. By passing through the QWP, the light is changed from horizontally polarized light into circularly polarized light. The circularly polarized light is normally incident on the mirror and travels back towards the QWP. This second pass through the QWP changes the circularly polarized

light into vertically polarized light. The vertically polarized light reflects at a 45° angle from the PBS and all the power that was originally aimed at the mirror is now directed to the right [figure 3-14] [ref 18].

Very similar techniques are used in the systems described in section III, although they are slightly more intricate because of the number of beam-combinations required. The optical systems described below were modulator-based, and thus required two types of beam-combinations. The first beam-combination involved the constant beam used to illuminate the transmitting modulators and the modulated data leaving the same node. The second beam-combination involved the modulated data from a preceding node reaching the detectors of the present node.

3.2.6) Image manipulation

The simplest interconnect to construct would be a set of uni-directional, point-to-

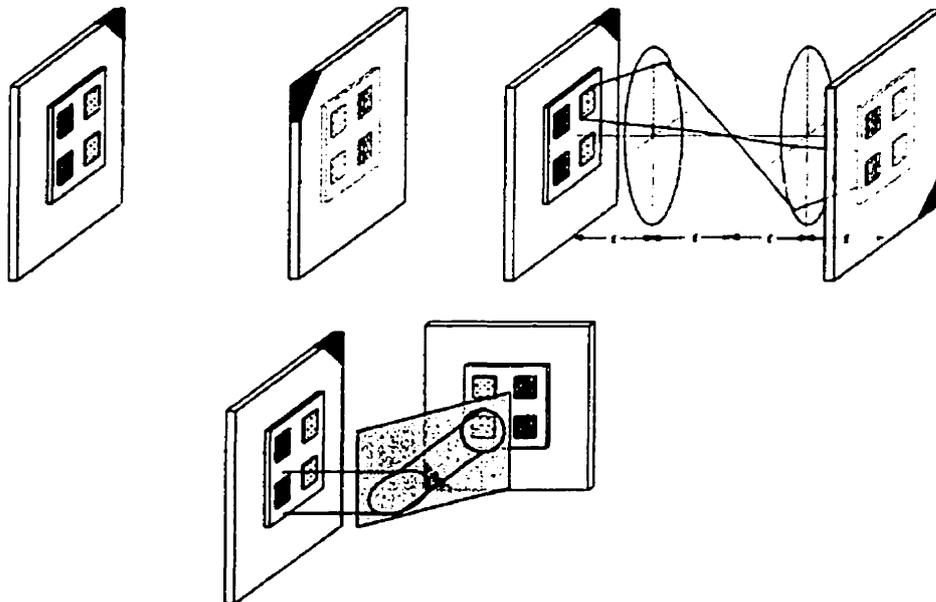


Figure 3-15: a) A direct path, b) an optical relay, c) a path off a mirror

point links with one chip transmitting the data and the other chip receiving the data. However, the design of the optoelectronic microchips (see Chapter 4 - VLSI Optoelectronics) required both optical transmitters and optical receivers on the same substrate. Therefore, a slightly more complicated design is needed because the

transmitters of one device must align with the receivers of the next device. For example, two optoelectronic microchips with optical transmitters on the left side and optical receivers on the right side can be arranged so they are facing each other, and the links will match up. Unfortunately, this requires enormous amounts of optical power for distances greater than a few millimeters, and is extremely susceptible to cross-talk. The solution is to use an optical relay to guide the light. A simple relay requires 2 lenses, of focal length f , and total separation of $4f$. The image in an infinite-conjugate relay is rotated 180° about the optical axis. For the transmitters (the light squares) to match up with the receivers (the darker squares), the second chip must also be rotated 180° about the optical axis. The reflection from a mirror is another mechanism that is sometimes used to change an image's orientation. In this scenario, the chip also needs to be rotated 180° about the optical axis for the transmitters to align with the receivers. However, this technique produces a slightly different transmitter-receiver mapping from the mapping obtained with the infinite-conjugate relay [figure 3-15].

These two methods of image manipulation can cause serious difficulties during a design. However, these same effects can also provide a means of re-orienting images so that they will map to the desired areas. In the design of the systems presented in the next section, both of these techniques were used.

3.3) System demonstrators

3.3.1) Introduction

The two demonstration systems presented in this section were designed, assembled, and characterized by members of McGill University's Department of Electrical and Computer Engineering - Photonic Systems Group during a span of roughly 5 years from 1995 to 1999. These systems were called the Phase II and Phase III demonstrator systems, and were constructed consecutively. Several people in the group were responsible for the design and implementation of the optical interconnects described below and they are gratefully acknowledged at the beginning of this thesis.

The following descriptions will draw on the numerous topics explored in the preceding sections; thus there will be little elaboration for most of the elements within

each system. The following descriptions of both systems will be highly qualitative in nature. The quantitative analysis of all the aspects of both the Phase II and Phase III designs have already been completed in papers by Y.S. Liu et al. and B. Robertson, respectively [ref 19, 9].

3.3.2) Phase-II optical system

The Phase II system was designed to be placed in a VME chassis [ref 20] [figure 3-16] and interfaced with multiple PCBs. The optical interconnect was to replace the electrical backplane of the VME bus structure. A typical VME chassis has up to 64-bits of parallel data, the optical interconnect was to have a 32 x 32 array of parallel optical transmitter-receiver links between each node. This would increase the number of interconnections between PCBs

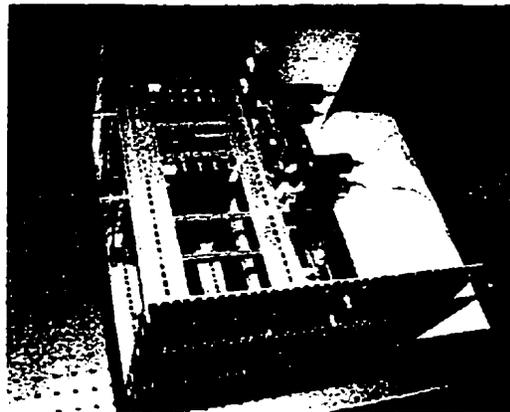


Figure 3-16: Photo of Phase-II system

by an order of magnitude. However, the Phase II prototype was only capable of supporting 4 PCBs with a 4 x 4 array of parallel optical transmitter-receiver links. Although, this was a proof-of-concept design using small optoelectronic chips, the full array could have been designed but at a much higher cost.

The optical interconnect linking the 4 PCBs was based on a closed-loop, uni-directional ring as discussed above. Each optoelectronic microchip was composed of a 4 x 4 smart pixel array capable of the modes of operation described previously (see Chapter 2 –Architecture). The receive state allowed optical data to be converted into electrical data. The transparent state allowed optical data to be re-transmitted as optical data. And the inject state allowed electrical data to be converted into optical data. The smart pixel array was designed using complementary metal-oxide-silicon (CMOS) processing circuits and multiple-quantum well (MQW) optoelectronic devices (see Chapter 4 - VLSI Optoelectronics). Since the MQW devices are modulator-based optoelectronics used in reflection mode, each node required combining three types of beams. The first type of

beam was a constant read-out beam focused on the transmitting modulators. The second type of beam was a reflected beam from the modulator and directed towards the next chip. And the third type of beam was the modulated beam from the previous chip onto the detecting receiver of the present chip

[figure 3-17]. The design of the optical receiver electronics

also required that each transmitter and receiver be composed of 2 MQW diodes for a total of 4 MQWs per smart pixel, this is called dual-rail encoding, and was done because of poor reflectivity from the MQW devices. Dual-rail encoding overcame the problem with providing a reference threshold voltage within the receiver circuit with which to judge ones and zeros (as explained in Chapter 4 – VLSI Optoelectronics). The beam

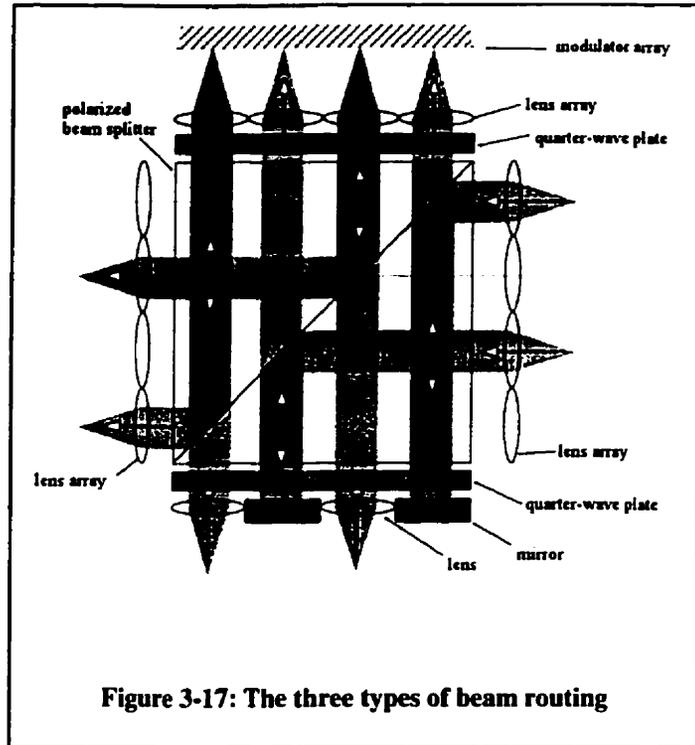


Figure 3-17: The three types of beam routing

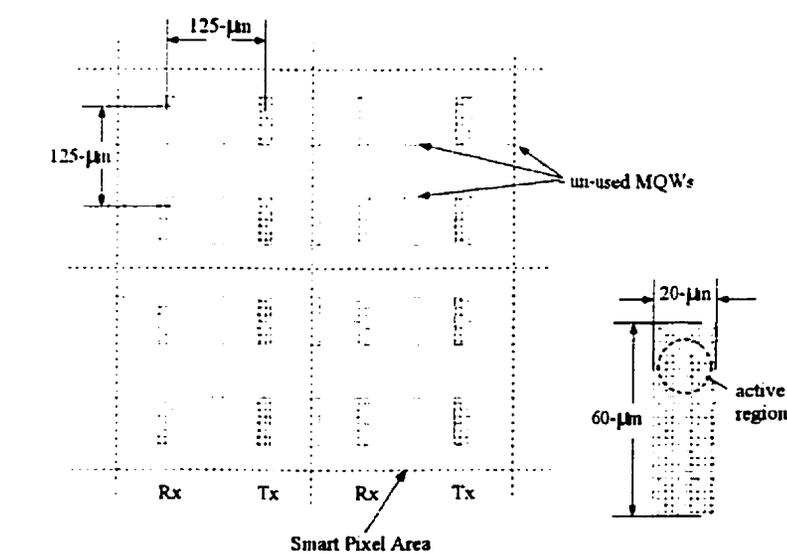


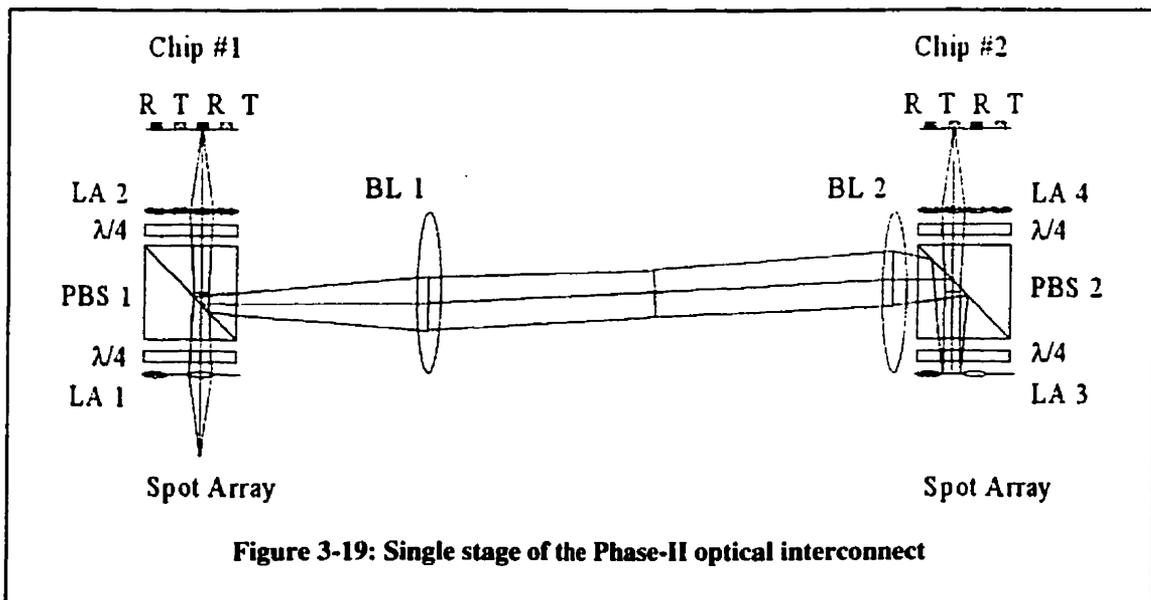
Figure 3-18: Layout of MQW device on Phase-II chip

combination module therefore had to route 32 constant “read-out” beams to the modulator devices, 32 modulated beams towards the next chip, and 32 beams arriving at the detector devices from the previous chip.

A total of 64 MQW devices were used

per node and they were patterned in a regular array with a pitch of 125- μm in both directions. The MQW devices were organized such that alternating columns corresponded to detector MQW diodes and modulator MQW diodes. This partitioning of receiver columns and transmitter columns is essential to the method of beam combination at each node, and will be described below. The active region of the MQW devices was roughly a 20- μm diameter area [figure 3-18].

Essentially, there were 6 types of optical elements used to construct the Phase-II relay between nodes: the micro-lens arrays (MA), the QWPs, the PBSs, the patterned mirror-microlens arrays (PMMA), and the bulk relay lenses (BL). There are several other components used to help with alignment, such as Risley prisms and tilt plates, as well as optics used to see into the system so that visual inspection was possible, but these elements are left out for clarity. The Phase II design was a hybrid optical system that used both diffractive MAs and refractive BLs. The MAs were used to guide beams on and off



the chip while the BLs were used to relay the modulated beams from one stage to the next. The BLs were required because the relay distance of the micro-lenses was too short to span the link between chips. Using the maximum lens-to-waist technique of Gaussian beam propagation described above for the microlens relay, the relay distance was calculated to be approximately 7-mm. Unfortunately, this distance did not provide a sufficient distance between adjacent chips. Therefore, an infinite-conjugate bulk relay

was introduced between chip modules to provide a larger working distance. The focal plane of the first lens (BL 1) was coincident with the middle of the first PBS (this is where the Gaussian spots achieved their minimum beam waist). It should be pointed-out that the BL relay was shifted towards the second node so that the focal plane of the second lens (BL 2) would exist at the middle of the PBS, but only after reflecting 45° off the PBS, passing through the QWP, reflecting from the PMMA and arriving at the middle of the PBS with the opposite polarization. This complicated path was required because the image relayed by the bulk lenses of the minimum waist had to appear as though it had traveled half the distance between a micro-lens relay. The focal length of each bulk lens was 35-mm and allowed the entire system to be stretched-out to 140-mm. In the figure [figure 3-19], a single stage of the Phase II optical interconnect is shown.

The BLs not only provided some working distance between the nodes of the interconnect, but they also provided the image inversion that allowed modulated data to arrive at detector MQWs. The image rotation of the BLs re-ordered the columns of the modulated data so that the first column was the last and the last was the first. This is the reason that the optoelectronic microchip was patterned with alternating columns of modulators and detectors. Without the image rotation of the bulk relay, the modulated beams would be directly mapped to their original positions on the next node.

Another interesting aspect of the design was how all 3-dimensions of space were

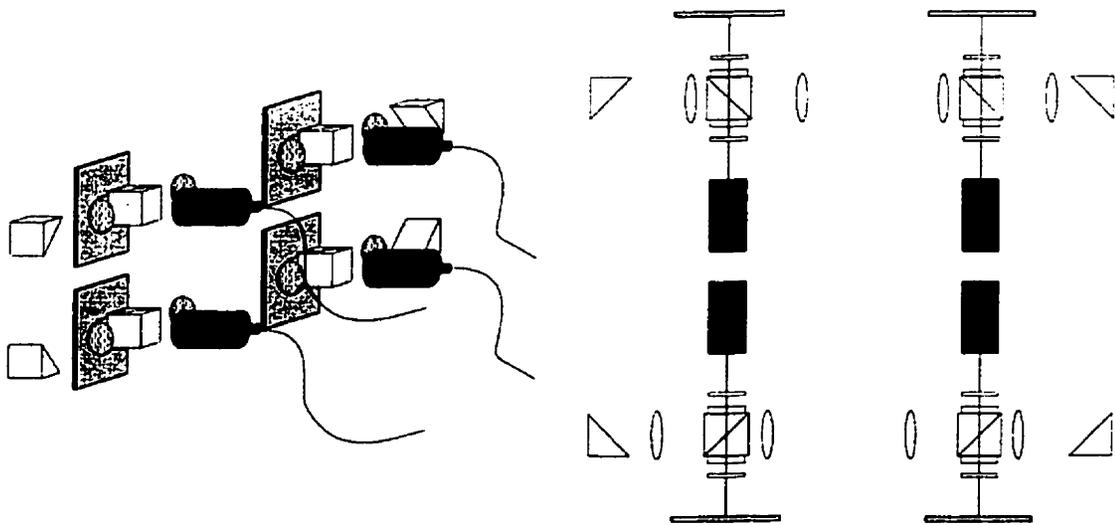


Figure 3-20: a) 3-dimensional concept of Phase-II optical system and b) flattened view of system

used to relay the beams [figure 3-20]. The optical relay transmitted modulated data in a closed-ring oriented in the y-z plane. However, the beams that were directed on and off the microchip were aligned with the x-axis with the surface of the microchip perpendicular to the direction of these incident beams. The constant read-out beams were also aligned with the x-axis. The closed-ring topology of the system was achieved by using 4 additional mirrors. The system could handle additional nodes simply by extending the number of relay links, but no more than 4 turning mirrors were required to close the ring in the y-z plane. Since 2 turning mirrors were used within the collimated path of the bulk relay, the image was unaltered, and the re-directed link appeared no

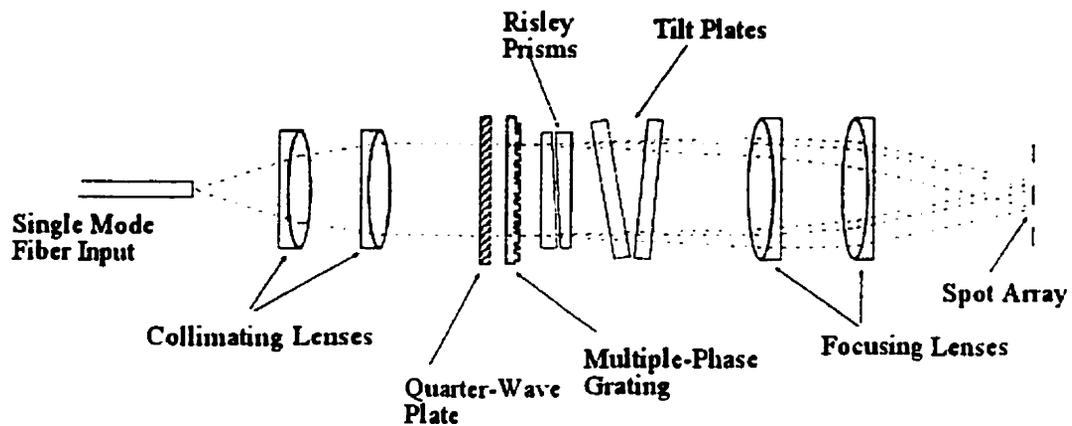


Figure 3-21: Optical spot array generator for Phase-II optical system

different than any other link.

A final aspect of the design concerns the source of light for the 32 spots incident on the modulator. The plane in the above figure labeled "spot array" had 32 constant intensity spots that were of minimum beam waist and expanding towards the first micro-lens array. These spots were generated using an optical power supply spot-array generator [figure 3-21] [ref 21]. The spot array generator used collimating and focusing optics as well as a computer generated

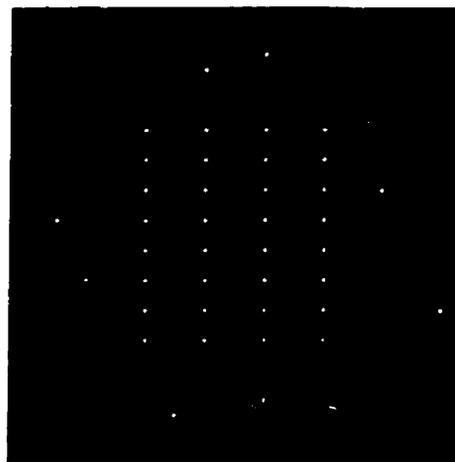


Figure 3-22: Spot array formed by generator

hologram in the form of a multiple level phase grating to generate multiple spots from a single input fiber. The optical power supply also had to ensure that the polarization of the spots was circular so that the desired effects in the rest of the system could be obtained. The multiple phase grating (MPG) is a glass substrate with multiple depth etch patterns. The MPG encodes phase delay information by way of different depths etched into the substrate. The pattern is an inverse spatial-Fourier transform of the desired spot array pattern. Therefore, by passing an incident plane wave through the MPG, many different order plane waves are created at different angles. As these various plane waves pass through the focusing lens, they are spatially transformed into the desired spot array pattern [figure 3-22]. This method allowed the entire optical system to be operated using a single laser source and a fiber-based delivery system.

3.3.3) Phase-III optical system

The Phase III system was the next generation of free-space interconnect for the optical backplane architecture. Although the architecture and chip technology remained similar to the Phase II system, the optical relay was significantly different. The optical design was still a uni-directional closed-ring interconnect, and each node was responsible for combining the same three types of beams, except in this case there were 1536 beams to combine at every stage. The design also took advantage of the 3-dimensional volume of space to interface with the chip as well as relay modulated data from one node to the next.

From the experience gained in developing the Phase II system, several aspects of the Phase III design were modified in very significant ways. The most drastic change was the size of the optoelectronic array on the microchip. The optical interconnect had to support a 16 x 16 smart pixel array requiring 1024 MQW diodes. The working area of the relay had to be increased from 4-mm² to 81-mm² to accommodate the larger MQW array. The aperture windows of the MQW devices were also increased from 20- μ m in diameter to 70- μ m in diameter. The size of the devices were increased because the Phase II system proved to be very susceptible to misalignment having such small optoelectronics. Another major modification was that the relay optics consisted of only telecentric mini-

lens arrays, the bulk relays did not exist. This, in turn, led to an alternate arrangement of MQW diodes on the surface of the microchip. Due to several beneficial characteristics described in another work [ref 22], the modulators and detectors were each grouped together to make an 8 x 8 array of 4 x 4 clusters. The columns of clusters were also arranged so the columns alternated between modulators and detectors. Detecting and modulating columns of 4 x 4 clusters were required to allow the modulated optical signals of one microchip to be aligned with the detector MQW diodes of the adjacent microchip.

Each 4 x 4 cluster used a mini-lens relay, thus the microchip required a total of 64 mini-lenses arranged in an 8 x 8 array. Each mini-lens was 800- μm in effective diameter and pitched 800- μm in either direction. This arrangement of mini-

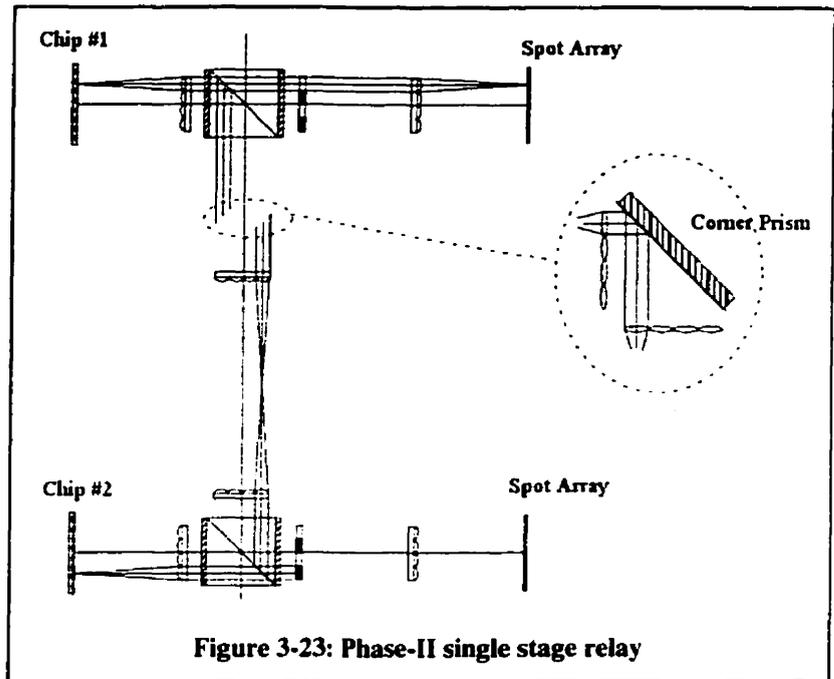


Figure 3-23: Phase-II single stage relay

lenses allowed a dense optical relay to be constructed, which had 2500 optical links/cm². While at the same time, it allowed a reasonably compact design. One of the more crucial aspects of the design was to balance the relay distance of the micro-lens arrays with the appropriate coverage of the clusters. Using a mini-lens diameter of 800- μm and a focal length of 8.5-mm, the Gaussian beam analysis predicted that 99% of the light from the cluster would be captured by the mini-lens. To increase the tolerance to misalignment, only infinite-conjugate relays were used. With an infinite-conjugate relay, the overall optical path length (OPL) from microchip-to-microchip was only 34-mm, which was not enough working distance for the microchips to be mounted into the system. The chips were already 8-mm x 8-mm leaving a 2.6-cm gap between them. From the electronic packaging aspect of the design, this was unacceptable. Therefore, a compromise was

reached and a 60-mm separation was targeted [figure 3-23]. Fortunately, the physical distance could be achieved as a result of the index of refraction of the optical elements. Even though the OPL was only 34-mm, most of the distance traveled by the beams was in glass, with an average effective index of refraction of 1.55. The path length was further extended to 58.8-mm by using an element with a dense index of refraction called the optical spacer, which was a co-planar rectangular block of glass.

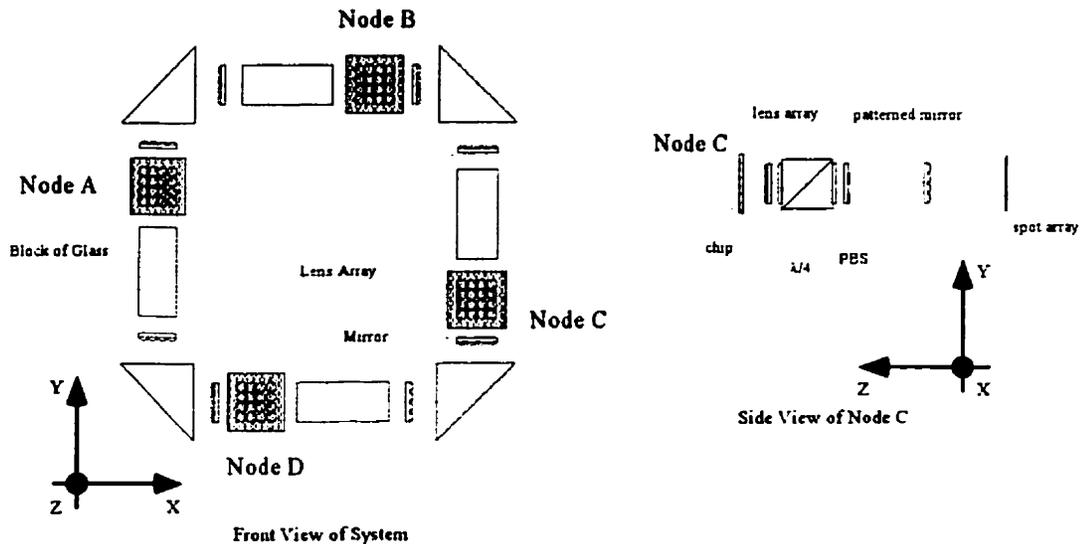


Figure 3-24: a) A front view of the Phase-III optical system, b) a side view of the optical system

The Phase III optical system was still based on the beam combination techniques of the Phase II design using patterned mirrors and PBS-QWPs, although the method of beam manipulation was quite different. In the Phase III system, each micro-lens relay maintained its modulated signals within the micro-lens channel. The re-ordering

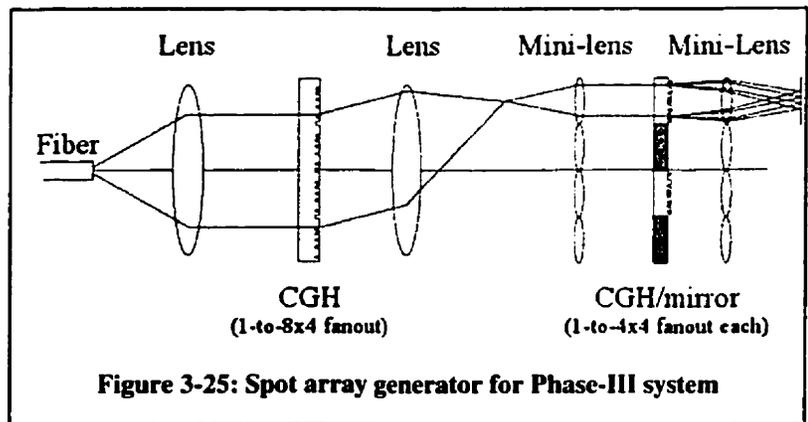


Figure 3-25: Spot array generator for Phase-III system

of the channels was obtained by using the reflection from the turning mirror. This mirror was located between each node and took the form of the corner prism, which were also used to close the ring [figure 3-24]. The spot array generator produced minimum waist

Gaussian spots at the spot array plane, these beams travel through the first micro-lens array and then through the patterned mirror. The patterned mirror was composed of vertical strips of reflective metal placed beside vertical strips of diffractive 1 x 16 fan-out gratings. These patterned mirrors implemented the second fan-out grating in the cascaded fan-out system. The first fan-out grating generated a set of 4 x 8 spots, each of these spots were then passed through the second fan-out grating creating a total of 512 spots incident on the modulator diodes [figure 3-25]. The beams were then directed towards the QWPs and PBS and then to the second micro-lens array. The beams were reflected from the modulators and the modulated data moved back through the micro-lens array and reflected at 45° from the PBS. The beams then traveled to the corner prism, which not only changed the direction of all the beams, but also re-ordered the columns of clusters. After the corner prism, the beams were focused and re-collimated and the double bounce

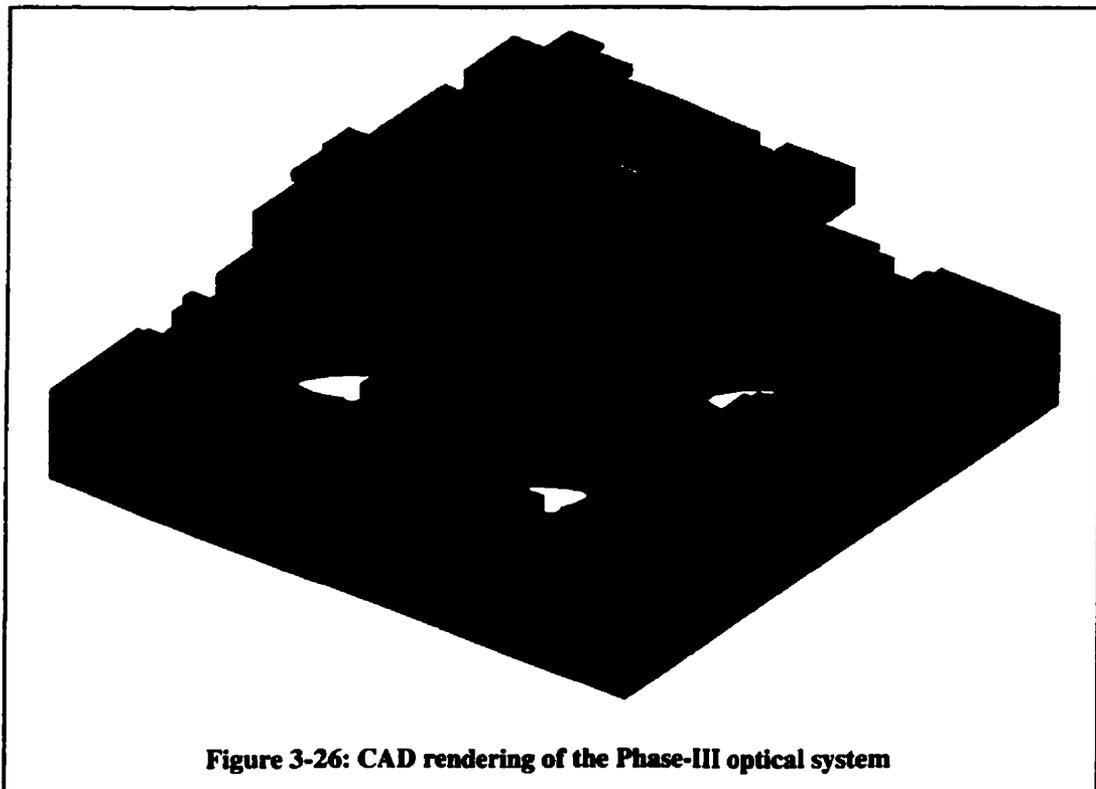


Figure 3-26: CAD rendering of the Phase-III optical system

procedure using the patterned mirror was used to direct the beams onto the detectors of the second microchip. A CAD rendering of the Phase III system is shown in [figure 3-26].

The final major difference between the two designs was its scalability. In the Phase II system, additional nodes could be inserted between existing nodes by increasing the overall length of the system. In the Phase III system, the general principal was the same, but the method by which additional nodes could be added was slightly less obvious. To increase the size of the system, the optical path would have to be arranged in a square-wave pattern. The corner prism, instead of turning back to close the ring, would be turned forward and the pattern could repeat in a forward progression [figure 3-27]. This staggered pattern was essential in order to provide a suitable pitch for the printed circuit boards that would be placed behind the chips. If a linear arrangement was adopted, the pitch between device planes, and similarly PCBs, would be between 8 and 10 cm. The square-wave pattern allowed tighter placement of PCBs, but they had to be placed at slightly odd vertical heights.

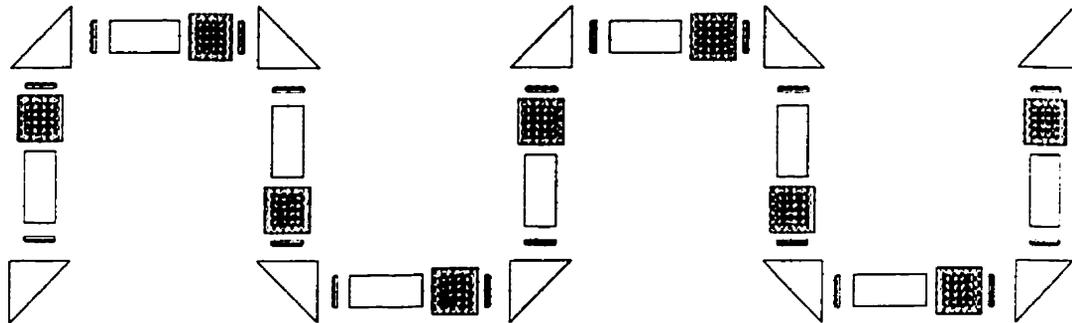


Figure 3-27: Scalability of the Phase-III optical system

3.4) Conclusion

This chapter has discussed the type of optical relays used by the microchips that will be described in chapter 4. The topics covered in this chapter discussed some of the methodology and some of the optical components used to create both the Phase II and Phase III optical systems. Each system was then outlined and certain aspects of their respective optical designs were discussed, especially when it concerned the interface between the optics and the optoelectronic microchips. The designs were not examined in great detail in this chapter, mainly because other papers and thesis explore these designs in depth. Even so, valuable insight into the design issues regarding the microchip were obtained. The difference between a regular array of optoelectronics on a microchip versus a clustered approach have a major impact on the complexity of the optical system, solving some issues in one aspect while introducing problems in another. The counterpart of this chapter is the microelectronic and optoelectronic design and analysis found in chapter 4. With a general knowledge of the optical issues involved in relaying signals from chip-to-chip, more care can be placed on the design and layout of the microelectronic interface.

3.5) References

- [1] A. Ghatak, K. Thyagarajan, Introduction to Fiber Optics, Cambridge University Press, New York, 1998.
- [2] J. M. Senior, Optical Fiber Communications: Principles and Practice 2nd Ed., Prentice-Hall International Series in Optoelectronics, New York, 1992.
- [3] F.A. Jenkins, H.E. White, Fundamentals of Optics 4th Ed., McGraw-Hill, New York, 1976
- [4] J.E. Midwinter, "Photonics in switching", IEE Proceedings-J, Vol. 139, No. 1, Feb. 1992, pp. 1-12
- [5] D. Kuhl, K. Drogemuller, et al., "PAROLI – a Parallel Optical Link with 15 Gbit/s Throughput in a 12-channel wide interconnection", The 6th International Conference on Parallel Interconnects (PI '99), 1999, pp. 187 -193
- [6] K.C. Pohlmann, The Compact Disc, Vol. 5, A-R Edition Inc., Madison Wisconsin, 1988
- [7] Optical design for photonics: summaries of papers presented at the Optical Design for Photonics Topical Meeting – Palm Springs, California, Optical Society of America, Washington DC, March 22-24, 1993
- [8] Y. Liu, Design, implementation and characterization of free-space optical interconnects for optical backplanes, Ph.D. Thesis, McGill University, Montreal, Canada, 1997.
- [9] B. Robertson, "Design of an optical interconnect for photonic backplane applications", Applied Optics, Vol. 37, No. 14, May 1998, pp. 2974-2984.
- [10] Centillion 1000 Multiservice ATM Solutions, Bay Networks Data Sheet (Nortel Networks), <www.nortelnetworks.com>
- [11] E. Hecht, Optics 3rd Ed., Chapter 4, Addison-Wesley, New York, 1998
- [12] C.D. Carey, D.R. Selviah, S.K. Lee, S.H. Song, J.E. Midwinter, "Computer-generated hologram etched in GaAs for optical interconnection of VLSI circuits", Electronics Letters Vol. 28, No. 22, Oct. 1992, pp. 2082 -2084
- [13] F.N. Borrelli, Microoptics technology : fabrication and applications of lens arrays and devices. Marcel Dekker, Inc., New York, 1999.
- [14] E. Hecht, Optics 3rd Ed., Chapter 5-6, Addison-Wesley, New York, 1998
- [15] J.T. Verdeyen, Laser Electronics 3rd Ed., Chapter 3, Prentice-Hall, New-Jersey, 1995
- [16] H. Kogelnik, T. Li, "Laser Beams and Resonators", Applied Optics, Vol. 5, No. 10, Oct. 1966, pp. 1550-1566.
- [17] F. Lacroix, M. Chateaneuf, X. Xue, A.G. Kirk, "Experimental and numerical analyses of misalignment tolerances in free-space optical interconnects", Applied Optics, Vol. 39, No. 5, Feb 2000, pp. 704-713
- [18] H.S. Hinton, An introduction to photonic switching fabrics, Plenum Press, New York, 1993.

[19] Y. Liu, B. Robertson, G.C. Boisset, M.H. Ayliffe, R. Iyer, D.V. Plant, "Design, implementation and characterization of a hybrid optical interconnect for a four-stage free-space optical backplane demonstrator", Applied Optics, Vol. 37, No. 14, May 1998, pp. 2895-2914.

[20] IEEE standard for a versatile backplane bus: VMEbus, ANSI/IEEE Std 1014-1987, 1987

[21] R. Iyer, Y.S. Liu, G.C. Boisset, D.J. Goodwill, M.H. Ayliffe, B. Robertson, W.M. Robertson, D. Kabal, F. Lacroix, D.V. Plant, "Design, implementation, and characterization of an optical power supply spot-array generator for a four-stage free-space optical backplane", Applied Optics, Vol. 36, No. 35, Dec. 1997, pp. 9230-9242.

[22] D.R. Rolston, B. Robertson, H.S. Hinton, D.V. Plant, "Analysis of a microchannel interconnect based on the clustering of smart pixel device windows", Applied Optics, Vol. 35, No. 8, March 1996, pp. 1220-1233.

Chapter 4: VLSI Optoelectronics

4.1) Introduction

The very-large-scale-integrated optoelectronic (VLSI-OE) microchip is a hybrid of technologies and attempts to solve the previously described problems (see Chapter 1 - Introduction) using light to transmit to and from the chip. The VLSI-OE chip merges electronic digital processing with the transmission capabilities of photonic technology [ref 1]. A VLSI-OE chip is usually composed of two components: the silicon digital processing circuits and a set of light sensitive optoelectronic devices. A VLSI-OE chip uses light-sensitive devices that have been patterned across its surface to communicate with similar chips. By using the 2-dimensional surface area of a chip to communicate, the bottleneck incurred by the previous electrical-only technique may be alleviated.

There are a total of four VLSI-OE chips that will be discussed in this chapter. The first two chips were part of a project called the Phase-II system. The next two chips were part of a project called the Phase-III system. Each iteration was a more aggressive design used to demonstrate that multiple optical beams of light could be relayed among multiple VLSI-OE chips. They were iterations of the same basic architecture presented in a previous chapter (see Chapter 2 – Architecture) increasing in either optoelectronic array size or complexity. The four chips were called: the Beta-Chip, the Workshop-Chip, the Phase-III-A chip and the Phase-III-B chip.

In this chapter, the four iterations of VLSI-OE chips will be briefly outlined. Specific parts of the architecture will be addressed for each chip. The general Floorplan and layout of the chips will also be given. The way the chip Floorplan and optical design influenced each other will also be discussed. Additional topics such as the optical receiver and transmitter designs will be covered, as well as details on the operation of the type of optoelectronic device used in the design.

4.2) The Beta-Chip

The Beta-Chip was the first VLSI-OE chip used in the McGill Photonic Systems group Phase-II demonstration system. The Beta-Chip was part of a multi-project chip that was part of a beta-site test-run for the Advanced Photonics Department of Lucent Technologies [ref 2].

4.2.1) Chip technology and optoelectronic specifications

The beta-site test-run multi-project chip was 3.7-mm x 3.7-mm in size and was partitioned into four quadrants of 1.7-mm x 1.7-mm each, one of which the McGill Photonic Systems group was able to use. The multi-project chip was fabricated using a p-substrate (n-well), 5-Volt, 3-metal layer, 0.8- μm gate-width CMOS process offered by Hewlett-Packard through the MOSIS foundry service. The transistor layout was done using Tanner's L-EditTM tools using the SCMOS (Scalable CMOS) design rules with a lambda of 0.5-microns. Prior to the design, Lucent determined certain design restrictions such as the limited use of the top-metal layer and a minimum gate width of 1.0- μm . This was done to ensure a high yield. The placement and the geometry of the optoelectronic devices was also specified by Lucent as well as several circuit designs, such as the optical

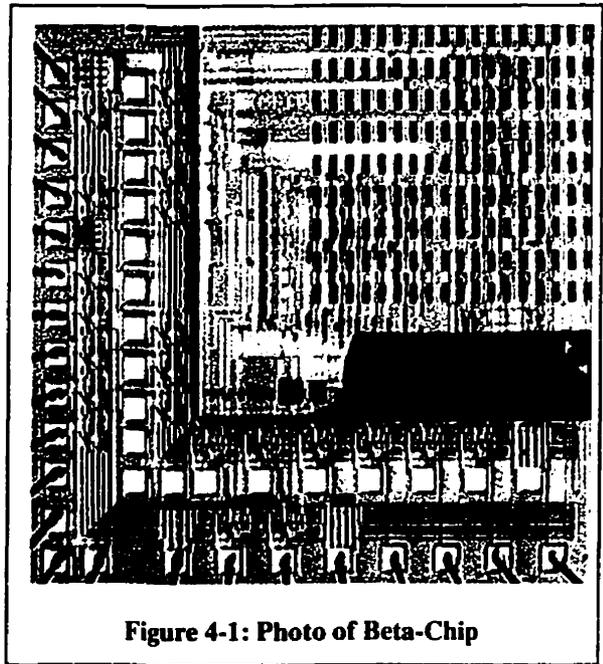


Figure 4-1: Photo of Beta-Chip

receivers [ref 3]. The Lucent group collated all the designs and had the multi-project chip fabricated. The multi-project chip was then post-processed at Lucent to attach the array of optoelectronic devices to the surface of the silicon chip. The optoelectronic devices were obtained from a 2.25-mm x 2.25-mm gallium arsenide (GaAs) substrate with a patterned array of rectangular (20- μm x 60- μm) multiple-quantum well (MQW) P-i-N diodes. The GaAs chip had a total of 18-rows and 36-columns of MQW devices with a center-to-

center column pitch of 62.5- μm and a center-to-center row pitch 125- μm . The silicon CMOS chip and the GaAs optoelectronic chip were attached through a process of flip-chip solder-bump bonding – the MQW structure and the flip-chip process will be described in a later section (see Section 4.6). The hybrid-chip was then diced into four quadrants leaving each quadrant with 8-rows and 16-columns of MQW diodes in one corner of each “sub”-chip (some of the MQW devices and the silicon along the quadrant cut-lines were sacrificed) [figure 4-1].

The 4 smart pixels on the Beta-Chip were arranged in a 2 x 2 array centered about the optical axis requiring a total of 16 MQW devices located near the inner corner of the diode array. This layout was necessary because the optical system used a 4- f system to relay the optical beams (see Chapter 3 - Optical Interconnects). A condition imposed by the optical design was that the MQW diodes had to be arranged as alternating columns of detectors and modulators. Since the micro-lenses used in the optical interconnect were 125- μm in diameter, an array of micro-lenses would perfectly match the pitch of the MQW devices in the vertical direction but would match only every second column in the horizontal direction. Therefore, the first and fifth columns were detector devices, and the third and seventh columns were modulator devices with the other columns left unused.

4.2.2) Digital design

The Beta-Chip smart pixel was based on a simplified version of the Hyperplane architecture (see Chapter 2 - Architecture). The Beta-Chip smart pixel was composed of an optical receiver, a bit-inversion multiplexer, a transmit multiplexer, an address comparison circuit, and an optical transmitter [figure 4-2]. There were a total of 61 transistors per smart pixel, and roughly 700 transistors for the entire design including the electrical bond-pad circuitry.

Each smart pixel used 4 MQW diodes to encode and receive optical data. Two MQW devices were used to transmit optical data and two MQW devices were used to receive optical data; this was called dual-rail encoding. The pair of transmitting MQW devices could encode a digital logic one as a high-low pair of beams and a digital logic zero as a low-high pair of beams. Dual-rail encoding doubled the number of devices and

the number of optical beams required. However, it simplified the receiver circuitry considerably.

The Beta-Chip smart pixel was an asynchronous circuit. It required no external clock signals. The address comparison circuitry was also asynchronous. A sufficient duration of time was determined

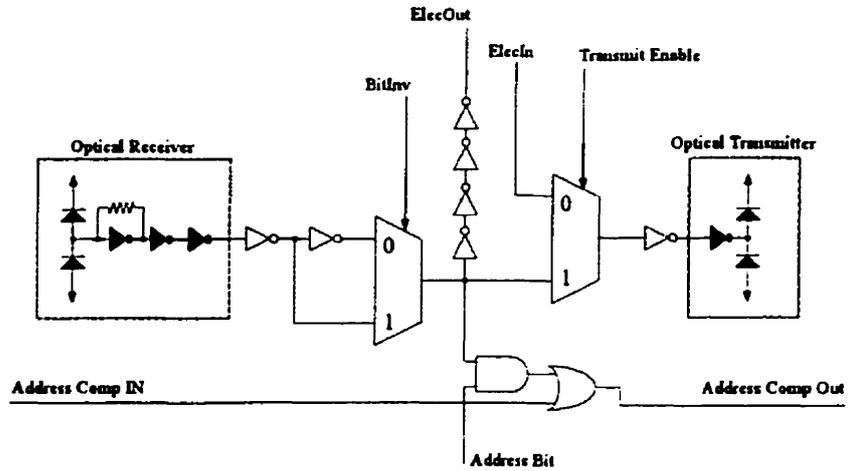


Figure 4-2: Beta-Chip smart pixel circuit

so that the output states of all address comparison circuits settled (including control signals) within the backplane before subsequent data was introduced. This was the basic technique for address comparison used throughout all the demonstrators, regardless of the complexity

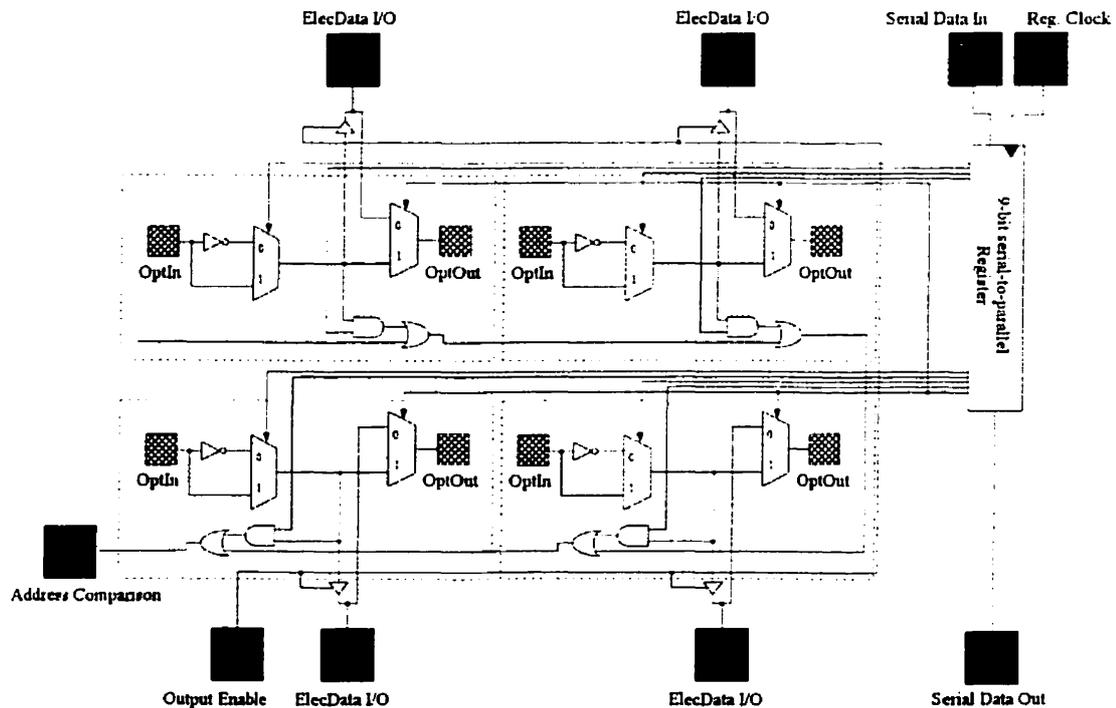


Figure 4-3: Complete Beta-Chip schematic

of the clocking or data paths.

Once the address comparison circuit had determined the desired state of the smart pixels in the channel, the state of the smart pixel was altered in one of three ways. If optical data was to pass through the smart pixel from optical input to optical output, the '*transparent state*' was used. Conversely, if data was to be optically received and electrically output, the '*receive state*' was used. Although normally the sender-reserve crossbar architecture required at least one channel per smart pixel array to remain permanently in the '*inject state*' (electrical to optical conversion), the Beta-Chip had only one channel, so the inject state was manually set. Due to a shortage of perimeter bond-pads, the electrical input and electrical output of each smart pixel were merged using bi-directional input/output bond-pads [figure 4-3]. This reduced the number of i/o pads from 8 to 4.

Since the electrical input and the electrical output shared the same bond-pads the '*extract state*' and the '*inject state*' were tightly coupled. By enabling the transmit

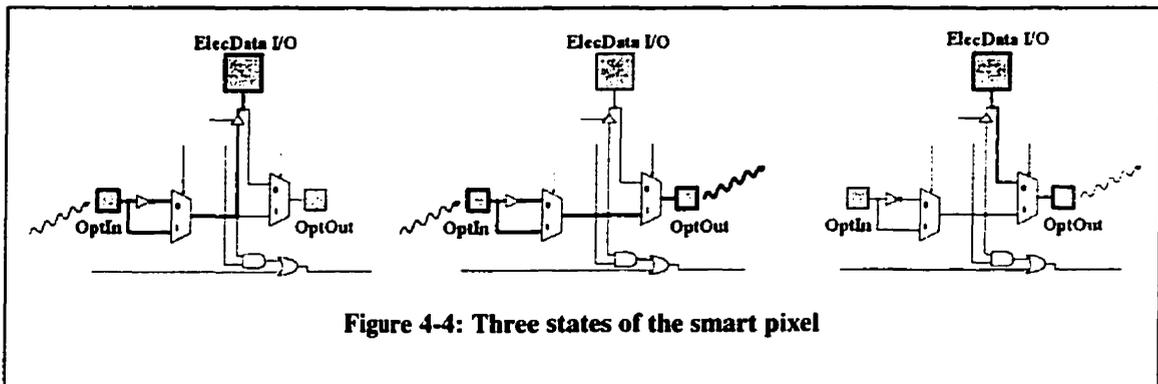


Figure 4-4: Three states of the smart pixel

multiplexer and setting the bi-directional i/o bond-pads in input-mode, the '*extract state*' was disabled and the '*inject-state*' was enabled allowing input electrical data to be routed to the optical output. When these settings were reversed, the transmit multiplexer received data from the optical receivers allowing for the '*transparent state*', and the bi-directional i/o bond-pads acted as output pads routing data off-chip from the optical receivers allowing the '*extract state*' [figure 4-4].

The function of the bit-inversion multiplexer was to complement the optically received data depending on a control bit stored in a serial register. The bit-inversion multiplexer could either pass the true signal or the complemented signal from the optical receiver. The need for such a mechanism was related to the nature of optical relay (see

Chapter 3 - Optical Interconnect). Due to the telecentric optical relay used between chips within the optical system, the dual-rail pairs of optical beams could become swapped such that a logical one issued by one chip could be received as a logical zero at the next chip. With the bit-inversion multiplexer, this problem could be avoided. The bit-inversion multiplexer was also the kernel for the ideas presented in the chapter on synchronization (see Chapter 6 - Synchronization).

The address recognition circuit was used to indicate when an incoming packet segment matched the permanent address assigned to the chip (see Chapter 2 - Architecture). The address recognition circuit linked all four smart pixels, and its output was directed off-chip. When the address recognition bond-pad was activated, a match had occurred. The address recognition circuit was the entity that defined the optical channel. Smart pixels that were linked by this addressing chain were part of the same optical channel. The address recognition circuit is given here [figure 4-5], and is based on a "one-hot-encoding" technique. Each smart pixel contains an identical functional block in the form of an AND-gate and an OR-gate.

The AND-gate could match an optically received logical high value with a logical high value in the permanent address. It would then pass this match along the chain of OR-gates. Using a "one-hot-encoding" technique, an address match would occur if at

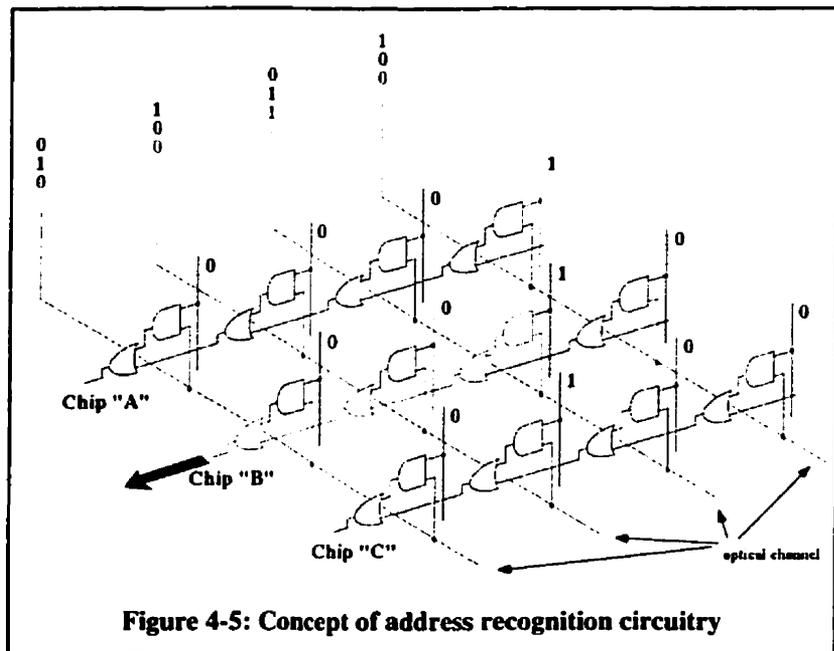


Figure 4-5: Concept of address recognition circuitry

least one optical data bit matched with a logical high value in the permanent address. The reason for this scheme was also related to the type of optical relay. Not only would the relay swap the pairs of beams, but it would also permute the order of the optical bits. The "one-hot-encoding" technique made it possible to uniquely identify each chip regardless

of the optical bit permutation. More on this topic will be presented in the next section (see Section 4.3.2).

The limited number of electrical input and output bond-pads was a very critical issue in the Beta-Chip design. The Beta-Chip smart pixel array had a total of 22 bond-pads, but only 6 pads were used to dynamically control and send data to the smart pixels. To overcome the need for additional bond-pads, a 9-bit serial-to-parallel shift register was used. The serial register was implemented using 3 bond-pads; the serial data input, the clock, and the serial data output. This register held the 4-bit fixed address of the chip used with the address comparison circuit, the 4 control signals for the bit-inversion multiplexers and a single bit, called the 'transmit bit', that was distributed to all 4 transmit multiplexers. This was a reduction from 9 static status bits to 3 bond-pads.

4.2.3) Layout

The area allotted to a single smart pixel was the first entity that had to be defined. This was ultimately a function of the MQW diode array. The Beta-Chip MQW diode was a 20- μm x 60- μm rectangle with 20- μm x 20- μm contact points at either end of the device. One contact point was the n-type side of the diode and the other was the p-type side of the diode. The silicon chip layout had to include an array of attachment points where the MQW array was to be placed. The attachment points were made from 25- μm x 25- μm top-metal pads with 18- μm x 18- μm openings in the top passivation layer (usually a coating of Silicon Nitride (SiN_x) is grown on the top surface of the chip during the last fabrication step to protect the circuitry below). The two attachment points of a single diode were separated by a center-to-center distance of 60- μm in the vertical direction. Each pair of attachment points had a vertical pitch of 125- μm and a horizontal pitch of 62.5- μm to match the GaAs diode array. Since the optical system required alternating columns of modulators and detectors, and because the pitch in both the horizontal and vertical directions had to match the square micro-lens array. Every second column of MQW devices was left unused. This provided a maximum total "repeatable" area of 250- μm x 250- μm for each smart pixel. Although there was a large amount of unused area that surrounded the four smart pixels, the smart pixel was restricted to the 250- μm x 250-

μm footprint so that techniques could be developed for future designs involving large arrays.

The bias lines for the totem-pole modulator pair and detector pair were laid-out in top-metal in wide, vertical lines between the attachment points. The biasing required 4 independent bond-pads to adjust the reverse bias across the MQW diodes to optimize the contrast ratio. The power and ground supplies were routed on the chip using wide metal-1 traces in long vertical lines. The smart pixel circuitry was composed of standard SCMOS library cells that were

37- μm high. The pMOS and nMOS transistors were equal size and roughly 12- μm wide with 1.0- μm gate lengths. The circuitry easily fit within the smart pixel area allotted [figure 4-6].

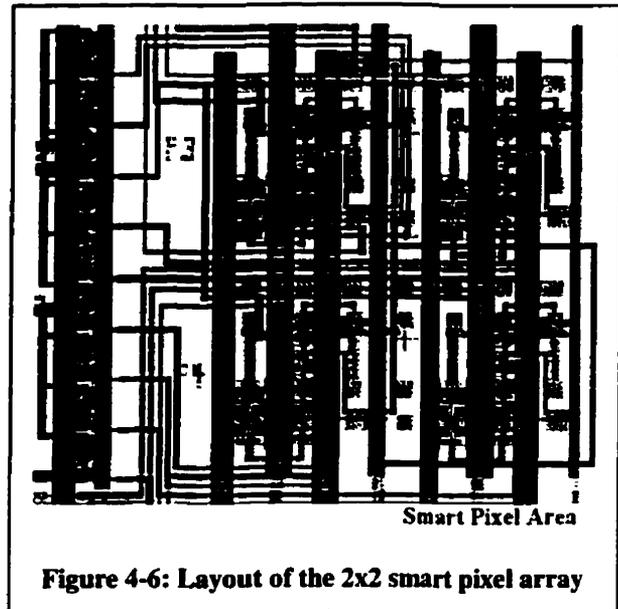


Figure 4-6: Layout of the 2x2 smart pixel array

4.3) The Workshop-Chip

The Workshop-Chip was the second iteration of VLSI-OE chip for the Phase-II demonstrator system. It was obtained through a CO-OP ARPA/Lucent hybrid SEED workshop that took place in 1996 at George Mason University (GMU) in Washington D.C.. The Workshop-Chip had a larger array than the Beta-Chip, but did not vary significantly from the basic design outlined in the Beta-Chip description.

4.3.1) Chip technology and optoelectronic specifications

The Workshop-Chip was a 1.95-mm x 1.95-mm hybrid VLSI-OE chip. It used the same process as the Beta-Chip; a p-substrate (n-well), 5-Volt, 3-metal layer, 0.8- μm gate-width CMOS process offered by Hewlett-Packard through the MOSIS foundry service. The transistor layout was done using Tanner's L-EditTM tools using the SCMOS

(Scalable CMOS) design rules with a lambda of 0.5-microns. The same layout restrictions as the Beta-Chip applied, such as the limited use of the top-metal layer and a minimum gate width of 1.0- μm . The same type of optoelectronic devices were also used. However, unlike the Beta-Chip, the Workshop-Chip was a complete chip where the MQW diode array was centered in the middle of the chip and there was access to all four sides of the chip for electrical bond-pads. The optoelectronic devices were obtained from a 1.6-mm x 1.6-mm gallium arsenide (GaAs) substrate with a patterned array of rectangular (20- μm x 60- μm) multiple-quantum well (MQW) P-i-N diodes. The GaAs chip had a total of 10-rows and 20-columns of MQW devices with a vertical pitch of 62.5- μm and a horizontal pitch 125- μm . The silicon digital chip and the GaAs optoelectronic chip were attached through a process of flip-chip solder-bump bonding.

4.3.2) Digital design

The Workshop-Chip smart pixel was almost identical to the Beta-Chip smart pixel. It contained an optical receiver, a bit-inversion multiplexer, an address comparison circuit, a transmit multiplexer, and an

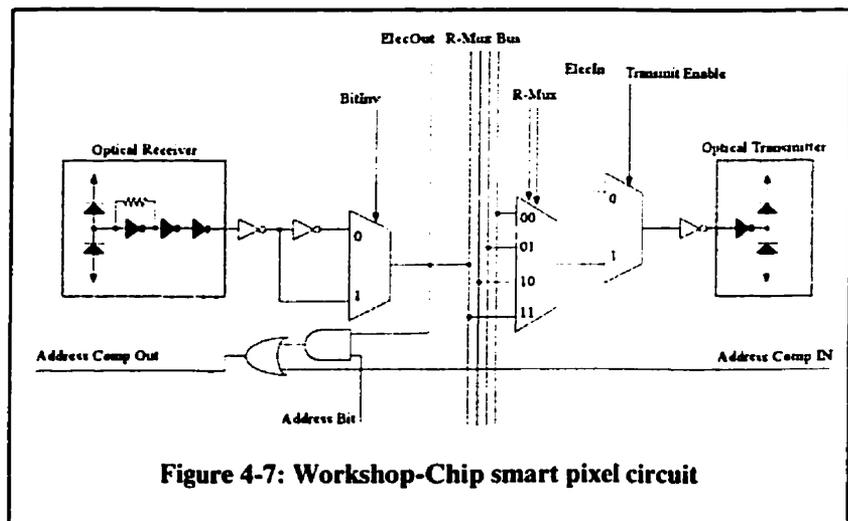


Figure 4-7: Workshop-Chip smart pixel circuit

optical transmitter. However, it contained one additional 4-to-1 multiplexer that could reconfigure optical channel propagation; it was called the R-Mux [figure 4-7].

The Workshop-Chip had a total of 16 smart pixels arranged in a 4 x 4 array for a total of 4 channels (rows) of 4-bits (smart pixels). Each smart pixel used 4 MQW diodes in dual-rail operation with the same optical receiver circuits and transmitter circuits as on the Beta-Chip. The Workshop-Chip smart pixel had approximately 99 transistors and a total of 3434 transistors for the entire design including the electrical bond-pad circuitry

and had a total of 44 bond-pads around its perimeter. Since the Phase-II optical system required that columns of detectors alternated with columns of modulators, half the MQW diodes in the array were not used. The first column contained detectors, the second was unused, the third column contained modulators, and the fourth was unused; this pattern continued across the chip. The array of small rectangles in the center region of the chip are the detectors and modulators [figure 4-8].

The Workshop-Chip smart pixel was completely asynchronous and could exercise the same three modes of operation as the Beta-Chip; the 'extract state', the 'transparent state' and the 'inject state'. Due to the number of smart pixels compared to the number of bond-pads around the perimeter of the Workshop-Chip, each smart pixel in the array could still be individually accessed. This allowed each channel to be controlled independently. However, the technique of merging the electrical

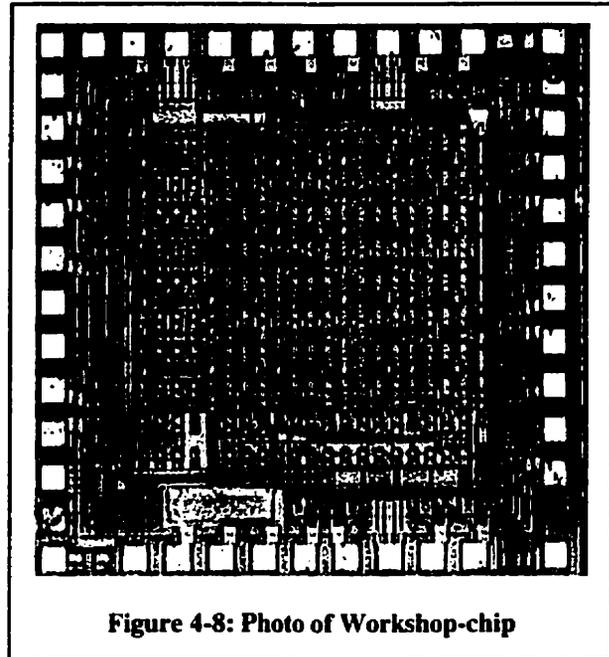


Figure 4-8: Photo of Workshop-chip

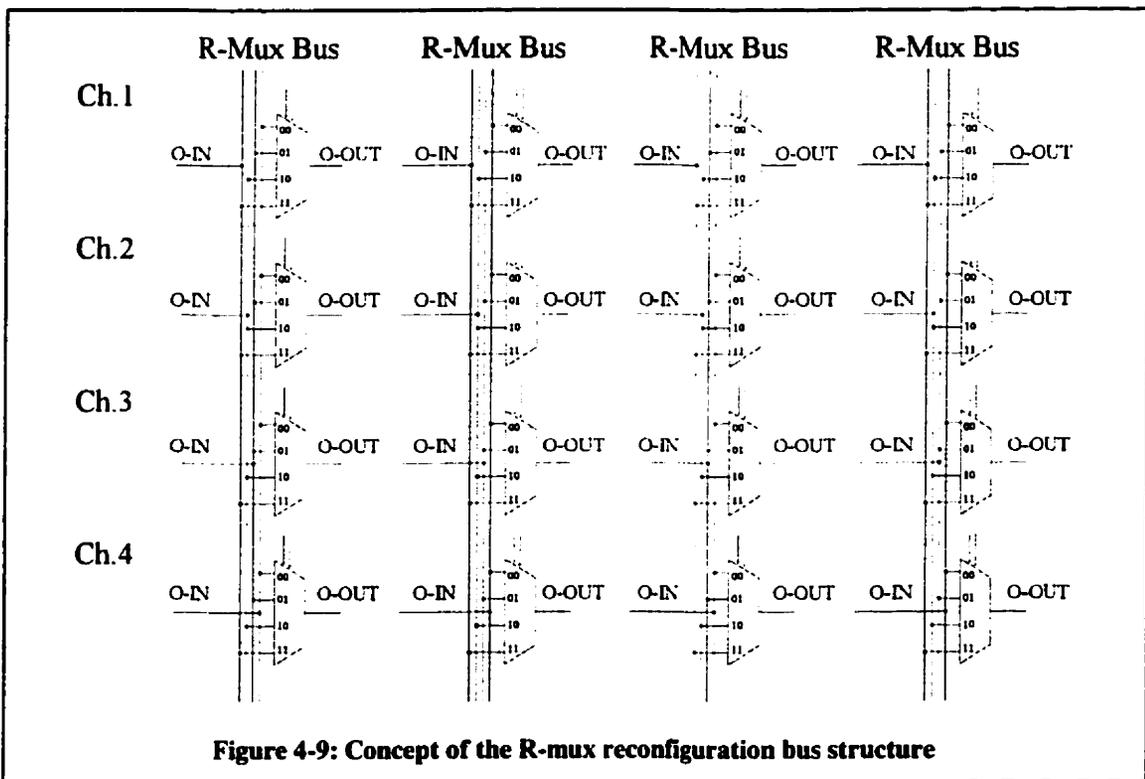
input and the electrical output of each smart pixel and routing it to a bi-directional bond-pad was still required. A single input bond-pad controlled the directionality of the bi-directional i/o bond-pads and it also controlled the state of the transmit multiplexer.

The Workshop-Chip had a serial-to-parallel register, similar to the Beta-Chip, used to store static information. The 16-bit serial register stored the permanent address of the chip, the states of the bit-inversion multiplexers and the state of the reconfiguration multiplexers on a per-channel basis, this ensured that all the smart pixels in a channel would behave in the same way. The serial register used 3 bond-pads; a serial data input, a register clock, and a serial data output (to check the data).

The reason for implementing the R-Mux was as a precaution against failures in MQW devices or optical paths. To ensure that a complete optical path around the entire system could be achieved, the R-Mux could electronically circumvent failed MQW

devices within a channel and continue propagating the data along other optical channels. If all MQW devices and all optical relay paths worked properly, the R-Mux would remain inactive, and the Workshop-Chip smart pixel would appear identical to the Beta-Chip smart pixel. There were 4 4-bit buses that spanned the 4 channels, each 4-bit R-Mux was connected to a particular bit-line. Optical data received by a channel could be re-transmitted on the same optical channel or re-routed to a channel above or below. If the state of the R-Muxes in the smart pixels were configured in a particular way, incoming optical data on one channel could be replicated and re-transmitted optically out multiple channels as well [figure 4-9].

Although the address comparison circuitry was identical to that in the Beta-Chip, the interpretation of the permanent address and the address header were considered in more depth in the Workshop-Chip design. The most important design feature of the



Hyperplane smart pixel architecture was the ability to partition an array of smart pixels into several unique parallel data paths (or channels). Each channel would be linked by a common address comparison circuit that propagated through each smart pixel in the channel.

The Phase-II system was designed to have 4 channels of 4-bits each. A sender-reserve crossbar embedding provided a unique channel to each chip in the optical system on which to transmit. To initialize a connection between two chips in the system, a particular chip on a particular channel would issue the 4-bit header segment. The header would circulate the ring and be compared with the permanent addresses in each chip in

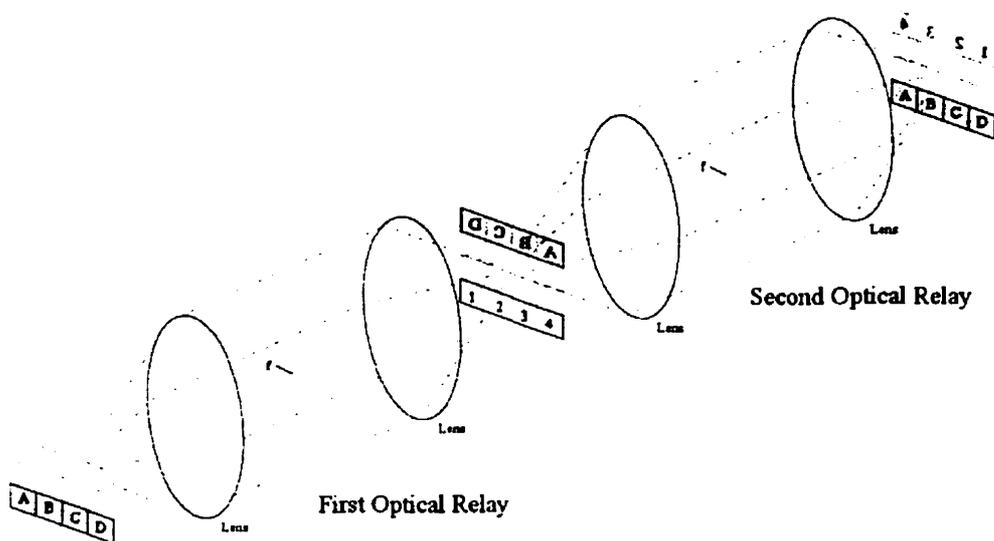


Figure 4-10: Image permutations through a set of infinite-conjugate relays.

the optical ring. If a match occurred, the subsequent data segments would be routed from the “sender” chip to the chip that matched (see Chapter 2 - Architecture). However, because of the telecentric optical relay between each pair of chips, any odd-number of “hops” around the optical system would cause the bits to be received in the opposite order. This was called “*optical bit permutation*”. A header segment that passed through an even number of chips would have also passed through an even number of telecentric optical relays resulting in no permutation of the bit pattern [figure 4-10]. For example, the path between chips #1 and #3, or between chips #2 and #4, has an even number of telecentric relays, thus no bit permutation would occur. However, if chip #1 sent a 0011 header segment corresponding to the address in chip #4, it would be received in the reverse order (i.e.: 1100) resulting in chip #4 not recognizing a match. Therefore, the addresses and the header segments had to be tailored to suit the optical system. The first alteration was to use a “*one-hot-encoding*” scheme for the permanent addresses in each

chip. A one-hot-encoded address has only one bit high and the others low in its address, such as: 0100. This required more bits to encode for the same number of chips in the optical system, but it helped uniquely identify nodes as well as lead to an easier mechanism for broadcasting data. For example, if chip #1 issued a 0101 header segment, it would match with both 0100 and 0001 allowing two chips to receive the data. Unfortunately, as indicated by the table [table 4.1], using one-hot-encoding was not sufficient. The header and the permanent address were structured in a certain way to allow for uniquely identified address matches. The permanent addresses on adjacent chips could not be identical upon bit reversal. The addresses: 0001, 0010, 0100, and 1000 were not allowed because the first and the last were identical upon bit reversal and so were the second and third. However, the addresses: 0001, 0010, 1000, and 0100 were acceptable. The table [table 4.2] shows a possible set of permanent addresses and header segments that were appropriate for a 4-chip optical system.

	Node #1 address (0001)	Node #2 address (0010)	Node #3 address (0100)	Node #4 address (1000)
Node #1 address (0001)	-	sent: 0010 received: 0100	sent: 0100 received: 0100	sent: 1000 received: 0001
Node #2 address (0010)	sent: 0001 received: 1000	-	sent: 0100 received: 0010	sent: 1000 received: 1000
Node #3 address (0100)	sent: 0001 received: 0001	sent: 0010 received: 0100	-	sent: 1000 received: 0001
Node #4 address (1000)	sent: 0001 received: 1000	sent: 0010 received: 0010	sent: 0100 received: 0010	-

Table 4.1: Address comparison with errors due to bit-permutations

	Node #1 address (0001)	Node #2 address (0010)	Node #3 address (1000)	Node #4 address (0100)
Node #1 address (0001)	-	sent: 0100 received: 0010 bit-inverted	sent: 1000 received: 1000	sent: 0010 received: 0100 bit-inverted
Node #2 address (0010)	sent: 1000 received: 0001 bit-inverted	-	sent: 0001 received: 1000 bit-inverted	sent: 0100 received: 0100
Node #3 address (1000)	sent: 0001 received: 0001	sent: 0100 received: 0010 bit-inverted	-	sent: 0010 received: 0100 bit-inverted
Node #4 address (0100)	sent: 1000 received: 0001 bit-inverted	sent: 0010 received: 0010	sent: 0001 received: 1000 bit-inverted	-

Table 4.2: Address comparison using appropriate address coding

The address comparison circuitry implemented on the Workshop-Chip was identical to the Beta-Chip, only replicated 4 times [figure 4-11]. This addressing scheme, in conjunction with the bit-inversion multiplexer (used to correct for individually

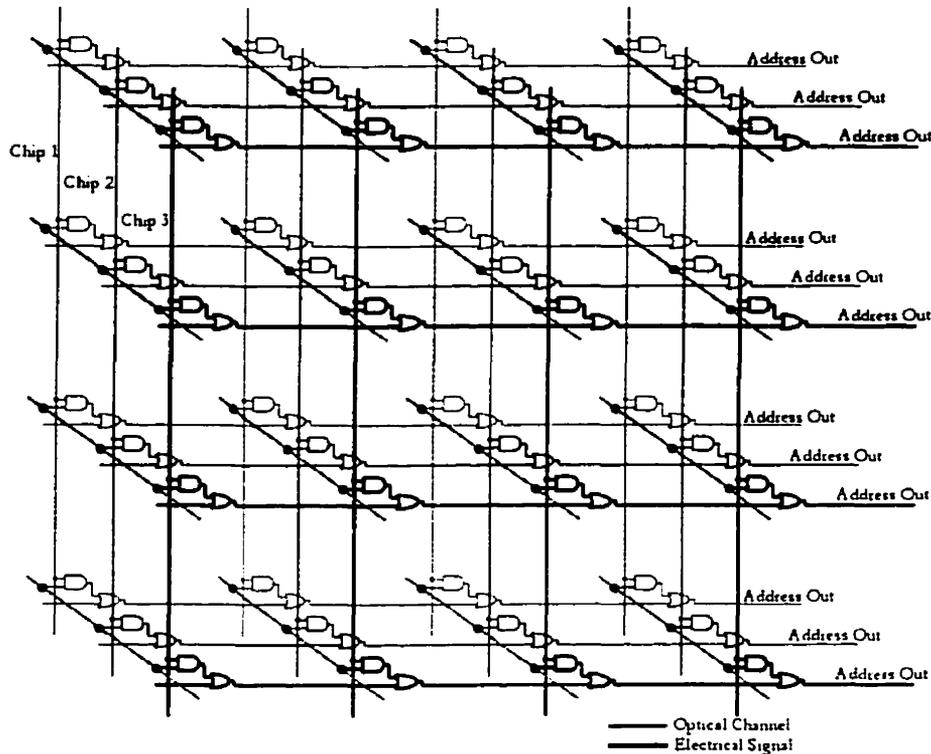


Figure 4-11: Representation of the Workshop-chip address comparison structure within the optical interconnect.

complemented optical bits), could be used to reverse the effects of both the “*optical bit inversion*” and the “*optical bit permutation*”. This could be done by interpreting the header segment. Although this was not implemented on any of the chips, fairly simple logic and a small amount of memory within the smart pixel could actively implement these functions.

4.3.3) Layout

The entire 10 x 20 MQW diode array was centered in the middle of the silicon chip, and the attachment points were identical to the Beta-Chip. All MQW diode attachment points were 25- μm x 25- μm squares of top-metal with 18- μm x 18- μm

passivation openings. All bias lines for the MQW diodes were supplied with wide, vertical top-metal traces running beside the attachment points. The size of the smart pixel was limited by the array structure of the MQW diodes. An area of 250- μm x 250- μm , similar to the Beta-Chip, determined the smart pixel dimensions and was the smallest replicable entity on the Workshop-Chip. This area constraint was more important than the Beta-Chip, because the smart pixels in the interior of the array were limited by the exterior smart pixels. The smart pixel layout was done using horizontally placed metal-1 trace lines for power and ground and vertically placed metal-2 trace lines for interconnects and bus lines.

The smart pixel circuitry was composed of standard SCMOS library cells that were 37- μm high. The pMOS and nMOS transistors were equally sized and roughly 12- μm wide with 1.0- μm gate lengths. To build a smart pixel cell that could abut on all four sides required that extra area within the smart pixel cell was allotted to the electrical data input and output for the entire channel. The data input and output lines as well as the control and address lines were routed horizontally through the array. Each channel had a reserved area below the smart pixel digital circuitry in which to route 11 lines into each channel (the 8 i/o lines and 3 control lines). The serial data register was arranged so that the first four bits of the register were placed at the top of each column of smart pixels. The following 12 bits were partitioned into 4 groups of 3-bits where each group was placed at the beginning of each channel

4.4) The Phase-III -A Chip

The Phase-III system was a more aggressive approach at implementing a large number of optical channels in a multiple board system. The Phase-III-A chip (P3A) was the result of a collaborative effort with several organizations and companies. However, unlike the Phase-II system design, virtually all the decisions concerning the general layout and Floorplan of the VLSI-OE chip were at the discretion of the McGill Photonic Systems group.

In the Phase-II VLSI-OE chip layouts, the size, the pitch, and the number of optoelectronic devices were dictated by Lucent. The Phase-II technology was originally

targeted towards fiber-based systems; the 62.5- μm and 125- μm device pitches were chosen to fit standard diameter optical fibers. The size of the active region for the optoelectronic devices was also chosen to be compatible with fiber-based systems; the 18- μm x 18- μm active area was suitable for most core diameters of multi-mode fibers. Unfortunately, the characteristics of the Phase-II VLSI-OE chip required a set of precisely aligned optics and opto-mechanical support structures to accommodate a very low tolerance to misalignment. However, the proposed optical system in the Phase-III project was a ground-up design that would ensure a less difficult assembly.

The basic VLSI-OE architecture was still based on a version of the Hyperplane architecture (see Chapter 2 - Architecture). However, the challenge with the Phase-III design was not only to demonstrate a multiple-board, high-speed optical interconnect, but also to demonstrate an enormous number of parallel optical connections within a compact volume. The total number of parallel connections was perhaps the largest incentive for researching the free-space optical interconnect.

4.4.1) Chip technology and optoelectronic specifications

The P3A chip was an 8-mm x 8-mm silicon chip and was fabricated using a p-substrate (n-well), 5-Volt, 3-metal layer, 0.8- μm gate-width, BiCMOS process capable of fabricating both MOS transistors and bipolar transistors. The silicon fabrication run was donated to the McGill Photonic Systems group as a fabrication grant from the Canadian Microelectronics Corporation (CMC) located at Queen's University in Kingston, Ontario. The fabrication run was performed by Nortel Semiconductors through the CMC. The transistor layout was done using Cadence CAD tools using a CMC supported BiCMOS design kit. The post-processing procedure was also under the control of the McGill Photonic Systems group. The optoelectronic devices were MQW P-i-N diodes used in reflection mode (see Section 4.6), but their pattern, size, and quantity could be tailored to the system requirements. The GaAs semiconductor and subsequent MBE growth was obtained from Dr. Anthony Springthorpe from Nortel Semiconductors. The device patterning was done in collaboration with Prof. John Currie and Dr. Edwis Richard (now in France), the GaAs chips were patterned at L'École Polytechnique (L'Université de

Montréal) LISA Lab. The final stage in this collaborative effort was the process of attaching the MQW diodes to the silicon chip and the subsequent GaAs substrate removal. This step was done by Dr. John Trezza of Sanders Corp. a division of Lockheed-Martin in Nashua, New-Hampshire [figure 4-12].

To increase the tolerance to misalignment of the optical system the active region of the MQW diodes was increased from that in the Phase-II system. Larger optical beams could be used, which in turn allowed for larger lenses and a more relaxed optical system. The modulating MQW diodes were made circular with an active region of 50- μm in diameter. The detecting MQW diodes were rectangular and had active regions of 70- μm x 70- μm . The modulator MQW diodes were made slightly smaller than the detector MQW diodes because the optical beams of light incident on the modulators would have traveled through less optics at that point and hence less aberrated. The smaller size could also reduce the capacitive load on the transmitter driver circuit, and thus achieve higher data rates. The detector was made larger and square to supply as large a target as possible to the modulated data beams, this would allow a greater misalignment tolerance. The large size of the detector MQW diodes did not limit the speed of the smart pixel because of the type of optical receiver circuit used (see Section 4.6.6.b). A total of 1024 MQW diodes were attached to the surface of the silicon chip using a technique called 'clustering' [ref 4]. The MQW diodes were arranged as an 8 x 8 array of 4 x 4 clusters. The pitch of the MQW diodes within a 4 x 4 cluster was 90- μm in both the horizontal and vertical directions. The pitch between clusters was 800- μm in both the horizontal and vertical directions, with an 800- μm diameter mini-lenses over each cluster. The mini-lens array was an 8 x 8 array of 800- μm diameter lenses. The size of the mini-lens allowed for an increased working distance and a more straightforward optical design. The clustering

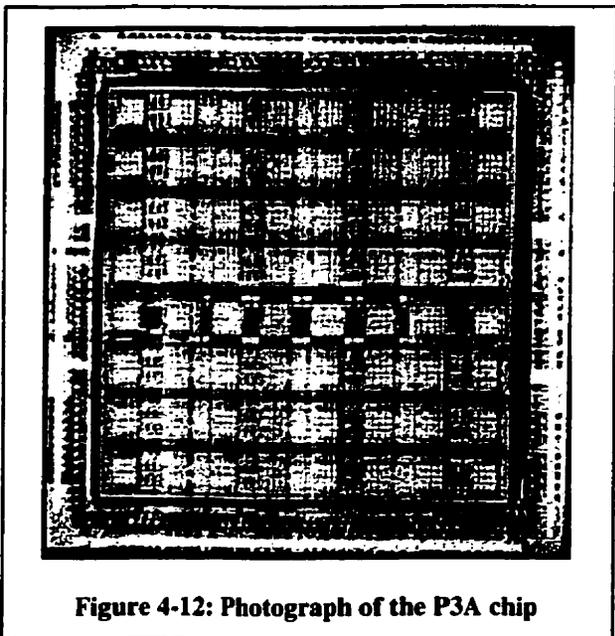


Figure 4-12: Photograph of the P3A chip

technique allowed the total area to remain small, while maintaining a high optoelectronic device density. Since the clustering technique allowed longer focal length lenses to be used, the optical relay could be designed as an infinite-conjugate system; a simpler optical design technique. The size of the GaAs chip that contained the clustered MQW diode array was 6.4-mm x 6.4-mm and was centered over the silicon chip and aligned with all the corresponding diode contact points.

The optical technique used to relay the light from a modulator to a detector required that columns of modulator clusters alternated with columns of detector clusters. The Phase-III optical system used a mirror to image the modulator array onto the detector array of the next chip. This required that the columns of MQW diode clusters alternate between modulator clusters and detector clusters, similar to the Phase-II optical system (see Chapter 3 - Optical Interconnect).

4.4.2) Digital design

Of the four chips, the Phase-III-A (P3A) chip had the most complicated smart pixel design. The P3A smart pixel [figure 4-13] was composed of several elements. A clocked-charge sense amplifier was used as the optical receiver (see Section 4.6.6.b), a testing multiplexer, a bit-inversion multiplexer, a primary D-FF, a secondary D-FF with an enable, an output concentrator multiplexer, a transmit multiplexer, an optical output

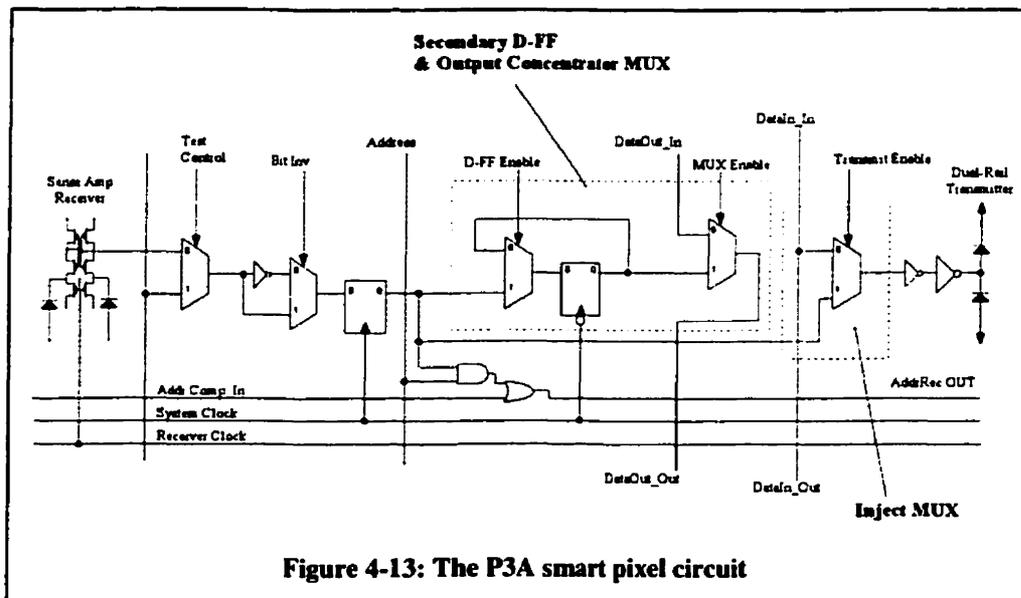


Figure 4-13: The P3A smart pixel circuit

driver, and an address comparison circuit. There were approximately 132 transistors per smart pixel and almost 40,000 transistors on the chip (including bond-pad circuitry). Each smart pixel was also supplied two high-speed synchronous clocks (HSS-Clocks) nominally out of phase by a quarter of a cycle. One clock was for the clocked-charge sense amplifier, and the other was for the primary and secondary D-FFs. These clocks also had to be distributed to all four nodes in the system in a highly synchronous manner (see Chapter 6 - Synchronization). The smart pixel also required three dynamically changing control lines called the D-FF enable line, the concentrator enable line, and the transmit enable line.

The method of passing data from one VLSI-OE chip to the next in the P3A design was similar to the way a pipeline architecture in most modern microprocessors. The method relied on the two high-speed synchronous clocks (HSS-Clocks) as well as the clocked-charge sense amplifier optical receivers and the primary D-FFs. For example, electrical data injected into an optical channel on chip A would travel to chip B via an optical data channel. The optical data would be received by the clocked-charge sense amplifiers and then latched by the primary D-FFs in chip B. The electrical data injected at chip A could now be changed, since chip B had sampled and stored the original data injected at chip A. The stored data at chip B could be passed to its own electrical output or to its optical output. If the data was re-transmitted optically to chip C, the data would be received by the clocked-charge sense amplifier and then latched by the primary D-FFs in chip C [figure 4-14]. Ideally, the process of progressing data along an optical channel could be done at a rate comparable to the internal operating speed of the microchip

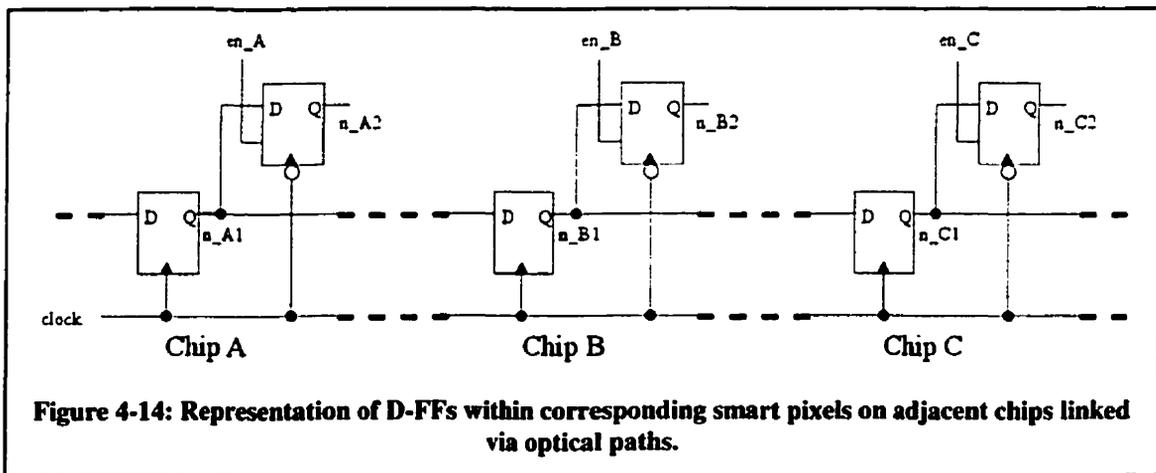
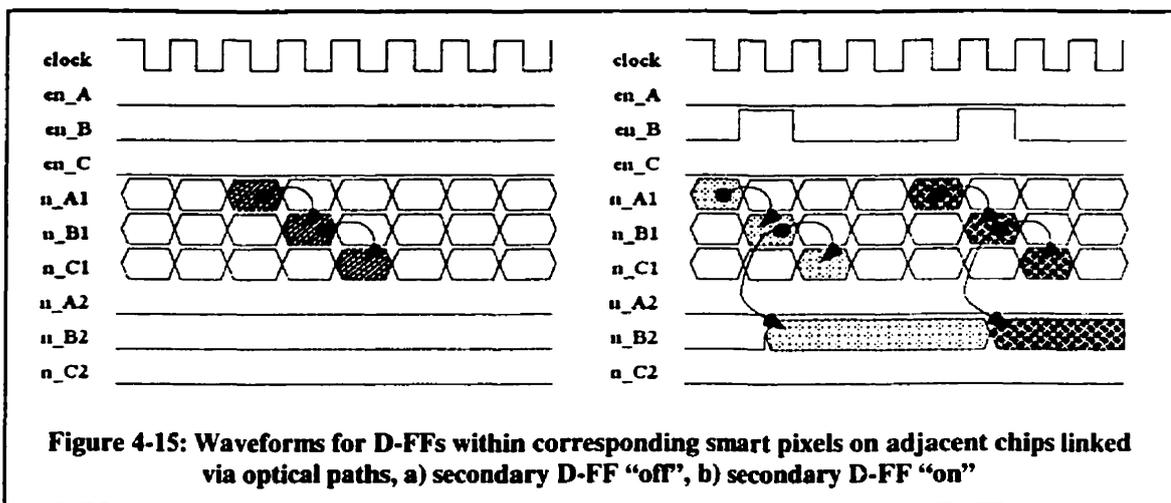


Figure 4-14: Representation of D-FFs within corresponding smart pixels on adjacent chips linked via optical paths.

(above 300-MHz for a 0.8- μm gate length CMOS technology) since the data never uses the electrical bond-pads. This technique also allowed for a form of time-division multiplexing (TDM) of the same optical channel. Since chip B stored the injected data from chip A, the injected data at chip A could be changed immediately after one clock cycle of the HSS-Clocks. Newly injected data could immediately follow the first, but be destined for another chip in the system [figure 4-15].

Although this technique seems to allow for very effective high-speed communications between distant VLSI-OE chips, the off-chip electrical data was not able to change at the same rate at which the data circulates in the optical system. Lower off-chip speeds have been the bottle-neck in many computer systems and are due to the



relatively large capacitive and resistive effects of metal trace lines when driving digital signals off-chip (see Chapter 1 - Introduction). Therefore, a second delay-element was used to facilitate the capture of off-chip electrical data. The output of the primary D-FF was routed to the secondary D-FF to "slow-down" the progression of data off-chip. The secondary D-FF would be enabled for the duration of one HSS-Clock period at the exact time when the appropriate optical data entered the smart pixel. The secondary D-FF would be immediately disabled and the data would be stored within the secondary D-FF for a longer duration of time. Meanwhile, subsequent optical data would be passing through the smart pixel from optical input to primary D-FF to optical output. The data stored in the secondary D-FF could now be directed to the electrical output bond-pads by enabling the concentrator multiplexer within the smart pixel. This would allow a longer

duration of time for the data to be driven off-chip [figure 4-16]. For a more complete description of this procedure, refer to the following reference [ref 5].

The injection and extraction structures of the P3A implementation were considerably different from the Phase-II system. Since there were 1024 MQW diodes and each smart pixel required 4 MQW diodes, there were a total of 256 smart pixels available on the chip. Each of the 256 smart pixels in the array required an electrical input and an electrical output. However, this would amount to a total of 512 bond-pads for electrical i/o data. Unfortunately, the size of the bond-pads in the design library limited the total number of bond-pads to less than 200. When the control, the MQW biasing and power supply bond-pads were subtracted from this total, there were less than 100 pads left for electrical i/o data. Therefore, the 256 smart pixels were grouped as 16 logical channels arranged in horizontal rows where each channel contained 16 smart pixels. The 16 channels were further divided into 2 groups of 8 channels, one called the upper-group and the other called the lower group. This was done to simplify access to the electrical i/o bond-pads. Each sub-group of 8 channels was allotted 16 electrical input data bond-pads and 16 electrical output data bond-pads, for a total of 64 electrical i/o data bond-pads around the perimeter of the chip

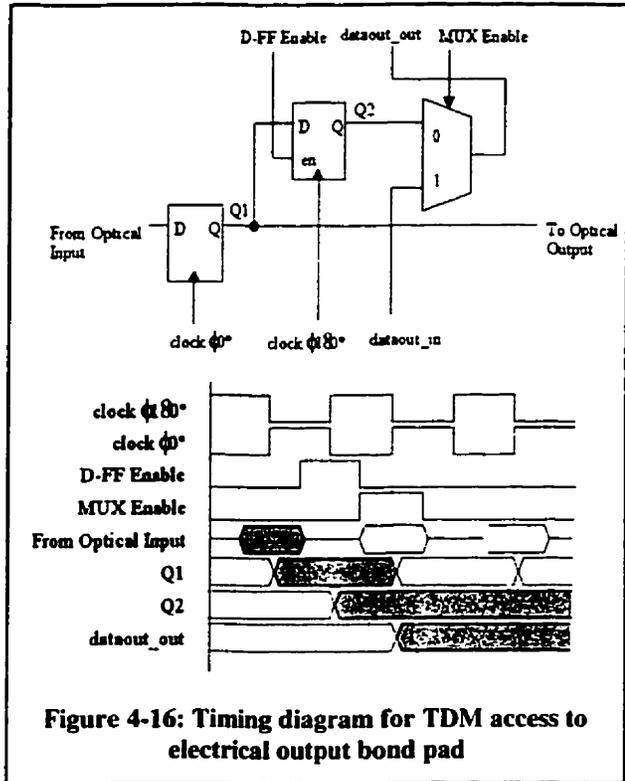


Figure 4-16: Timing diagram for TDM access to electrical output bond pad

Floorplan of the P3A/P3B chips. The floorplan shows a 16x16 grid of smart pixels. The columns are labeled 1 through 16 at the bottom, and the rows are labeled 1 through 16 on the left. The columns are grouped into four pairs of 4, labeled R, T, R, T at the bottom. The rows are grouped into two groups of 8, labeled CH 1-8 and CH 9-16 on the left.

Figure 4-17: Floorplan of the P3A/P3B chips

[figure 4-17]. By sub-dividing the 16 channels into 2 banks of 8 channels a larger data-path was provided (i.e.: 2 16-bit input ports and 2 16-bit output ports). In addition, there was a decrease in the access time to the deepest channel (instead of 16 channels deep, there were 2 banks of 8 channels deep).

To access the channels, two structures were used to transmit and receive electrical

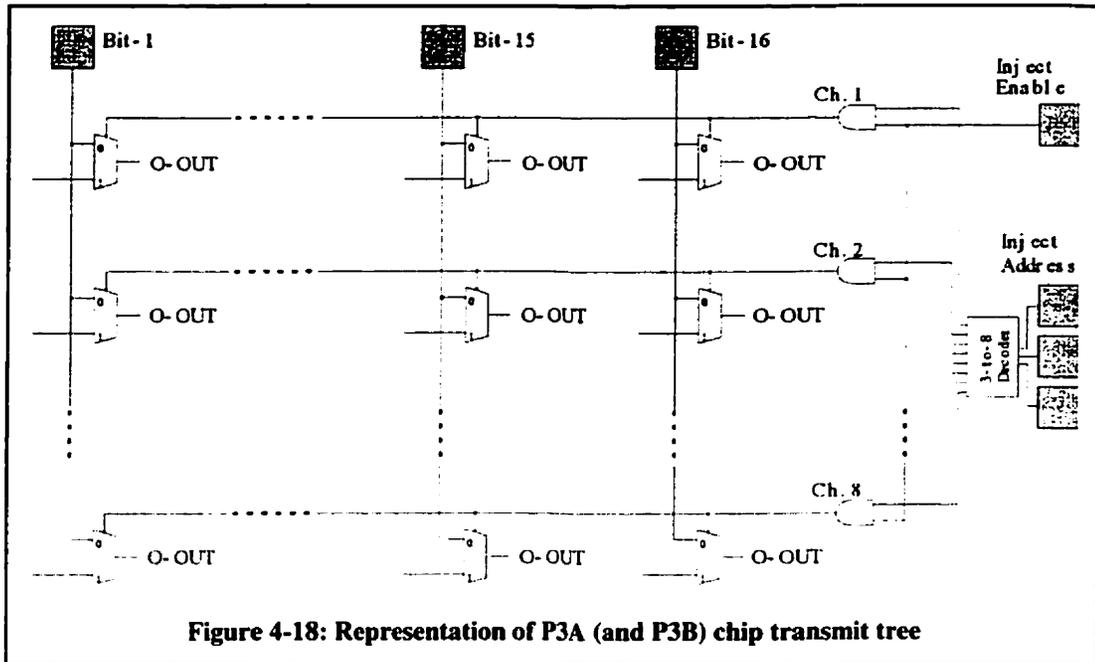


Figure 4-18: Representation of P3A (and P3B) chip transmit tree

data; these were called the input transmit-tree and the output concentrator. There were two input transmit-trees and two output concentrators, one pair was for the upper-group of 8 channels and the other pair was for the lower-group. The input transmit-trees were simple bus structures 16-bits wide and distributed to their respective 8 channels [figure 4-18]. The sender-reserve architecture allowed the transmit structure to be less complicated than the output concentrator because the reserved-transmitter channel was set during system start-up. Since there were 2 groups of 8 channels, two 3-to-8 channel decoders, requiring a total of 6 bond-pads, were used to select the correct transmitter channels. An additional pair of bond-pads (one for the upper-group, one for the lower-group) were distributed to the transmit multiplexers within the selected transmit channels, and were used to enable a channel at the correct time to inject the data into the optical data path.

The output concentrator consisted of cascaded 2-to-1 multiplexers, 16-bits wide, and passing through the 8 channels to the electrical output port [Figure 4-19]. Once a

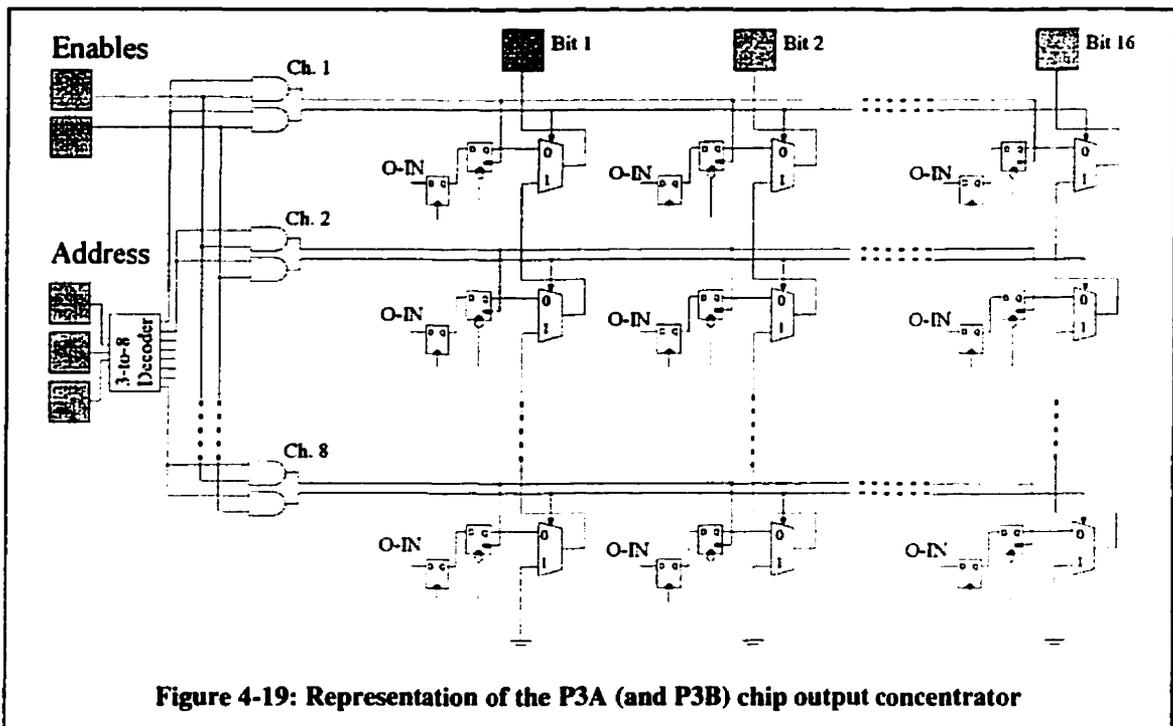


Figure 4-19: Representation of the P3A (and P3B) chip output concentrator

channel had been selected for receiving data, all the channels lower in the concentrator would be blocked from accessing the electrical output port. This was the reason for the *partial* crossbar interconnect architecture because the chip could only receive data on one channel at a time. This scheme also allowed for a natural channel priority as well. The channel closest to the electrical output port would have the highest priority and could override lower channels. The two output concentrators required a total of 6 bond-pads for the two 3-to-8 channel decoders and an additional two bond-pads (one for the upper-group, one for the lower-group) were used to enable the secondary D-FFs within a selected channel. These control lines were used more dynamically than the control lines of the input transmit-tree because they had to interact with the address recognition circuitry in each channel.

The address recognition circuitry in each channel functioned in the same manner described earlier (see Section 4.3.2) except it had 16 smart pixels. Each channel had its own address recognition output bond-pad (8 for the upper-group and 8 for the lower-group). When any of these address recognition circuits signaled an address match,

external processing circuitry was designed to process the match and return the appropriate 3-bit address to the output concentrator's 3-to-8 decode circuit. The external circuitry was also designed to enable the channel at the correct times to extract the data from the optical channel. This process was based on header segment information and timing from synchronized external clocks (see Chapter 2 - Architecture).

There were 2 28-bit serial-to-parallel registers that required a total of 6 bond-pads. One register was for the upper-group and the other was for the lower-group. A single register contained 8-bits for the test multiplexers, 4-bits for the bit-inversion multiplexers, and 16-bits for the permanent address. The test-multiplexers were used to introduce a static zero or one to simulate the operation of the clocked-charge sense amplifier optical receiver. This allowed testing of the digital part of the smart pixel before attempting to operate the optical receiver. The test multiplexers were present in every smart pixel. However, the state and the value of the test-bits were common to all the smart pixels within a column. The bit-inversion multiplexers were similar to those found in the previous designs but were distributed in the same way as the testing multiplexer. The control signals for the test multiplexer and the bit-inversion multiplexer were common to each column of super-cluster (described next).

There were a total of 189 bond-pads used on the P3A chip. Of the 189 pads, 21 pads were part of the advanced smart pixel design (which is not covered in this thesis [ref 6]), 7 pads were part of a temperature-sensor array used to monitor the temperature of the surface of the chip (which is not covered in this thesis [ref 7]), 22 pads were power, 22 pads were ground, and 12 pads were used to bias the MQW diodes. The remaining 105 pads were for the digital access to the smart pixel array.

4.4.3) Layout

The clustered array of MQW diodes patterned on the surface of the silicon chip required a different smart pixel layout than the Phase-II system. In the Phase-II system, each smart pixel could be patterned as an array, where the boundaries of the smart pixel were a function of the repeating MQW diode array. This strategy was very convenient for VLSI layout because a repeatable cell could be constructed. Unfortunately, the fine pitch

and small active regions of the MQW diodes made the optical system considerably more difficult to assemble. On the other extreme, if the VLSI-OE chip had been patterned with all the modulators on one side of the chip and all the detectors on the other side of the chip, the optical design would have been far easier to implement. However, the VLSI layout strategy would have required a significantly different approach because a repeatable cell would be more difficult to define. The clustering technique was an attempt at balancing the optical design requirements with the VLSI chip layout requirements while still maintaining a high connection density.

In the clustered approach, the optical design again required alternating vertical columns of detecting and modulating clusters. Therefore, the smallest region that was “perfectly” repeatable and contained both modulators and detectors was a horizontal rectangle with an area of $1600\text{-}\mu\text{m} \times 800\text{-}\mu\text{m}$ covering both types of cluster. This region was called a “super-cluster” and contained 16 modulator devices and 16 detector devices

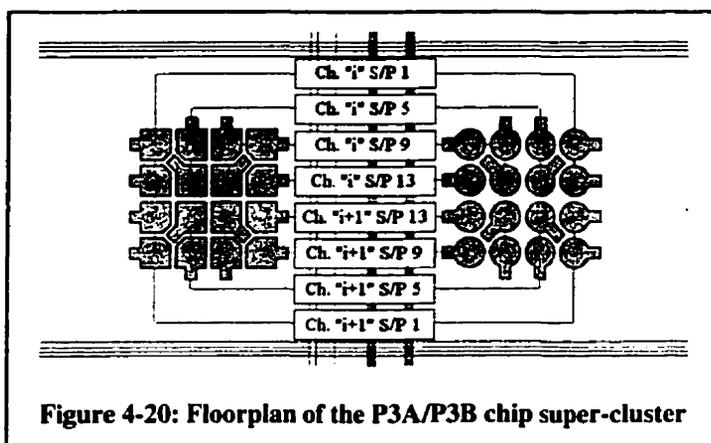


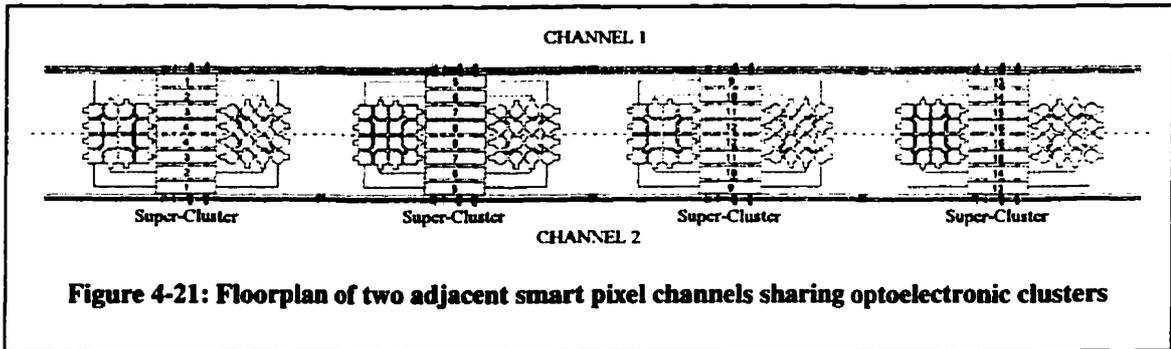
Figure 4-20: Floorplan of the P3A/P3B chip super-cluster

allowing a total of 8 smart pixels to be formed (assuming dual-rail MQW diode structures) [figure 4-20]. Although the super-cluster was a repeatable structure, it did not define the organization of the channels. The 16 channels were defined as horizontal rows stacked vertically on top of one another. This allowed the most freedom when accessing the electrical bond-pads along the top and bottom of the chip. When 4 super-clusters were placed horizontally side-by-side, they contributed 16 smart pixels to the first channel and 16 smart pixels to the second channel.

The super-cluster was partitioned between two unique channels through its middle horizontal axis [figure 4-21]. The upper-group of 8 channels consisted of 4 columns and 4 rows of super-clusters and accessed the top set of bond-pads. The lower-group of 8 channels also consisted of 4 columns and 4 rows of super-clusters but accessed the bottom set of bond-pads. The super-cluster had to abut properly on all sides because data

and control lines had to be routed vertically through the super-cluster, and address matching and channel enable lines had to be routed horizontally.

The major constraint placed on the layout of the super-cluster was to limit the amount of circuitry placed below the MQW diode clusters (especially when only two metal layers were available). Circuitry below a cluster could be susceptible to scattered



light that had missed the MQW diodes. The stray light would be absorbed by the silicon (penetration depth in silicon at 852-nm is about 15- μm) and generate minority carriers that could diffuse into the transistor circuits causing abnormal bit error rates or raising the optical power required to receive an optical bit. There was also the possibility of generating CMOS latch-up conditions within the parasitic bipolar transistors below the surface of the CMOS gates [ref 8]. Therefore, the region allotted to the smart pixel layout was confined to an area 800- μm high by roughly 450- μm wide centered between the detector and modulator clusters. The entire P3A chip layout was custom designed, where each smart pixel was only 38- μm high and roughly 450- μm wide. The control lines, the power, the ground, and the MQW bias lines were routed horizontally through the super-cluster, above and below the 8 smart pixels. Each smart pixel also had to be vertically symmetric because they had to abut once stacked on top of each other, this was done so that power or ground supply lines could be shared between adjacent smart pixels. All smart pixel transistors were sized with a balancing width ratio ($W_P = 2W_N$ for 0.8-micron BiCMOS) and a minimum transistor width of 1.8- μm to allow equal rise and fall times. The smart pixel used metal-1 for horizontal routing, and some metal-2 was used within the smart pixel in special horizontal tracks, but primarily it was reserved for vertical routing. Four electrical input lines and 4 electrical output lines were routed vertically

through each of the 4 columns of super-cluster for a total of 16 inputs and 16 outputs on each side of the chip.

The contact points for the MQW diode clusters on the silicon chip were made using top-metal with passivation openings. Each contact point was a top-metal square that was 30- μm x 30- μm in size with a passivation opening of 20- μm x 15- μm centered on the metal square. The corresponding contact on the MQW diodes were 15-mm x 20-mm and orthogonal to the openings on the silicon to obtain a better alignment during the flip-chipping process. Since each MQW required two contact points, there were 32 contact points per cluster. The modulator cluster required 8-pairs of totem-pole MQW diodes with a voltage, $V_{\text{mod+}}$, bias line and another voltage, $V_{\text{mod-}}$, bias line. The detector cluster used differentially connected pairs of MQW diodes for its sense-amplifier receiver circuits and required only one voltage, V_{det} , as a bias line (see Section 4.6.6.b). Both the modulator and detector clusters were laid-out symmetrically about the horizontal middle axis of the clusters. Symmetry was very important due to the nature of the optical design. The optical relay would rotate the image through 180° between the modulator plane and the detector plane. Therefore, the modulator totem-pole arrangement had to map to a similar detector arrangement so that pairs of dual-rail signals would not be split between two smart pixels.

The final characteristic of the Phase-III optical design was that the VLSI-OE chip had to be rigidly fixed to the first mini-lens array in a package called the chip module [figure 4-22]. This technique was used to increase the tolerance of several

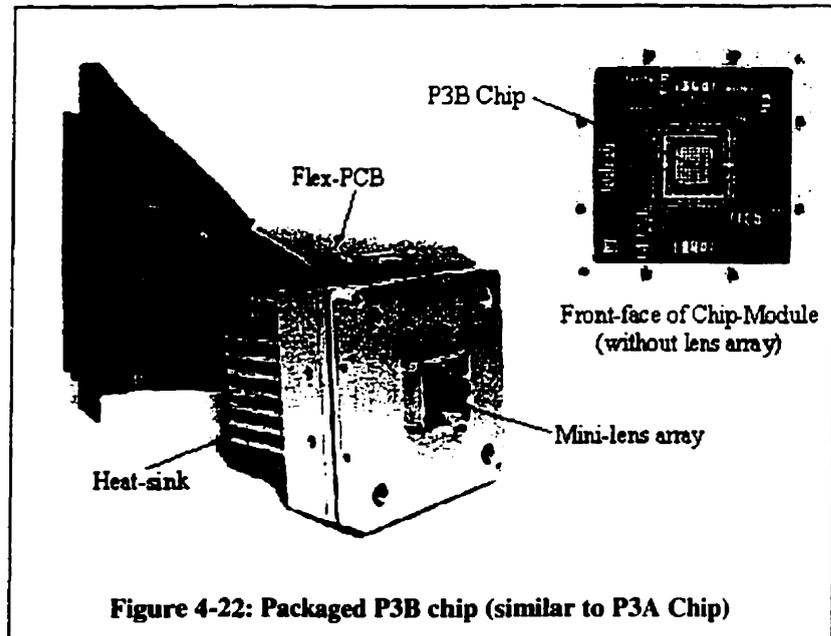


Figure 4-22: Packaged P3B chip (similar to P3A Chip)

types of optical misalignment. However, this required additional features on the chip that

could aid with the construction of the module. Therefore, structures such as quadrant-photodetectors, Fresnel interference mirrors, alignment markers and Talbot structures were placed on the chip. These topics are covered in detail elsewhere [ref 9,10].

4.5) The Phase-III -B Chip

The Phase-III-B (P3B) VLSI-OE chip was built because of several design and manufacturing errors with the Phase-III-A chip. The Phase-III optical system had not been exercised with an operating chip due to these errors and thus a second chip was required. The chip technology chosen for this iteration was less advanced than the technology used in the fabrication of the P3A chip. The minimum feature size on the P3B chip was twice as big as that on the P3A chip, resulting in a chip that would operate much slower. The reason for this choice was primarily due to the high cost of implementing the design in a smaller line-width technology. However, the primary objective was to demonstrate a *fully functioning*, massively parallel optical interconnect. Although obtaining high data rates was an issue, it was decided that it was more important to demonstrate the enormous connection density of the optical interconnect.

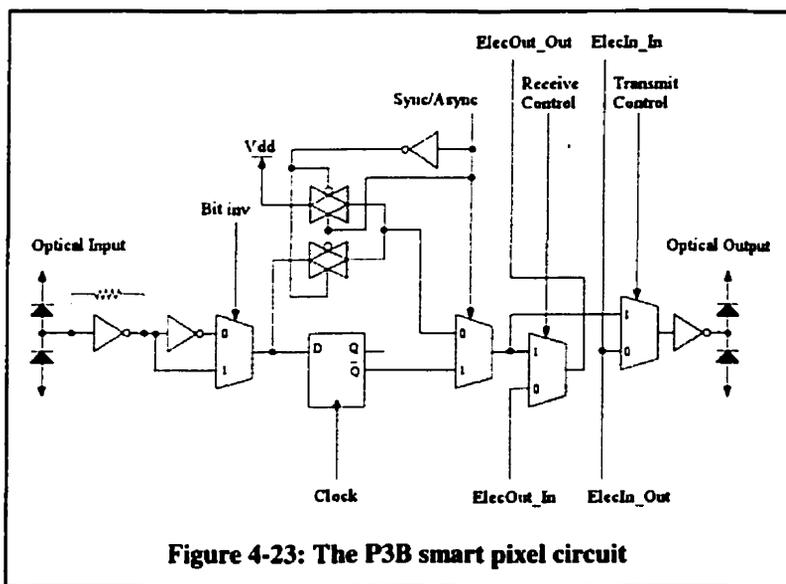
4.5.1) Chip technology and optoelectronic specifications

The P3B chip was a 9-mm x 9-mm silicon chip and was fabricated using an n-substrate (p-well), 5-Volt, 2-metal layer, 1.5- μm gate-width, CMOS process. Due to the n-doped substrate, the chip's substrate was required to maintain a +5-Volt bias. The silicon fabrication run was donated to the McGill Photonic Systems group as a fabrication grant from the Canadian Microelectronics Corporation (CMC). The fabrication run was performed by Mitel Semiconductors Corp. through the CMC. The transistor layout was done using Cadence CAD tools using a CMC supported Mitel15 CMOS design kit. The optoelectronic devices were MQW diodes used in reflection mode and the pattern of the diodes was under the control of the McGill Photonic Systems group. However, Dr. John Trezza of Sanders Corp. did the fabrication and the flip-chip attachment of the MQW diodes.

The pattern of the MQW diodes on the P3B chip was very similar to that on the P3A chip. Again, 1024 MQW diodes were attached to the surface of the silicon chip in an 8 x 8 array of 4 x 4 clusters. The clusters were pitched at 800- μm in both directions, and the diodes within each cluster were pitched 90- μm in both directions. The modulator diodes were again 50- μm in diameter and the detector diodes were 65- μm x 65- μm rectangles. The mini-lens array was identical to that described for the P3A chip; an 8 x 8 array of 800- μm diameter lenses each centered on top of a cluster. The chip was also organized such that columns of modulator clusters alternated with columns of detector clusters.

4.5.2) Digital design

Although the Phase-III-B (P3B) chip was slightly less complicated than the P3A chip, more test structures and more flexibility in the digital design was incorporated into the chip. The P3B smart pixel consisted of a transimpedance amplifier optical receiver, a bit-inversion multiplexer, a D-FF, two pass-gates and a by-pass multiplexer, an output



concentrator multiplexer, a transmit multiplexer, and a transmitter driver [figure 4-23]. There were 93 transistors per smart pixel and almost 30,000 transistors on the chip (including bond-pad circuitry).

Two key simplifications to this design were the absence of the clocked-charge sense amplifier optical receiver and the secondary D-FF. These components were left out because of the difficulty in providing several high-speed synchronous clocks to multiple chips in the system as well as the relatively poor performance of the clocked-charge sense amplifier (see Section 4.6.6.b). To simplify the testing of the chip, each smart pixel was also provided a mechanism to

run synchronously or asynchronously. During asynchronous operation, an external clock was not required. To allow asynchronous operation and by-pass the synchronous D-FF, a 2-to-1 multiplexer, based on MOSFET pass-gates, was used. The by-pass circuit was structured so that during synchronous operation, a meta-stable state from the optical receiver would not affect the rest of the chip through the sync./async. multiplexer's "0" port. During synchronous operation, the same type of pipelined operation outlined in the P3A chip would occur without the need for the secondary D-FF.

The same number of MQW diodes were used (1024) and each smart pixel required 4 MQW diodes, for a total of 256 smart pixels. The 256 smart pixels were grouped as 16 logical channels of 16 smart pixels arranged in horizontal rows. The 16 channels were divided into 2 groups of 8 channels, one called the upper-group and the other called the lower group and each sub-group of 8 channels had 16 electrical input data bond-pads and 16 electrical output data bond-pads, for a total 64 electrical i/o data bond-pads.

The injection and extraction structures of the P3B implementation were identical to the P3A structures [figure 4-18] [figure 4-19]. Two input transmit-trees and two output concentrators were used; one pair was for the upper-group of 8 channels and the other pair was for the lower-group of 8 channels. The input transmit-trees were simple bus structures 16-bits wide and distributed to their respective 8 channels. However, to simplify the design, each channel was supplied its own transmit bond-pad which could individually select a channel for transmission. Furthermore, the first channel in each of the 2 groups of 8 channels were sub-divided into 4 4-bit nibbles which could transmit independently from each other. A total of 22 bond-pads were used to control the transmit states (11 bond-pads for each group of 8 channels; 4 bond-pads for channel #1 and 7 bond-pads for channels #2 to #8). Unlike the 3-to-8 decoder circuits that could only decode a single channel at a time, the P3B implementation allowed multiple channels to be selected in a broadcasting manner.

The two output concentrators were also similar to the P3A implementation. They consisted of cascaded 2-to-1 multiplexers, 16-bits wide, that passed through each of the 2 groups of 8 channels. Just as the P3A chip, a channel priority existed, channels lower in the concentrator path would be blocked from accessing the electrical output port. Hence

the chip was still a *partial* crossbar interconnect architecture. To simplify the design, the output concentrators were also supplied individual channel select bond-pads. Similar to the transmit path, the first channels in each of the 2 groups of 8 channels were subdivided into 4 4-bit nibbles which could be placed in the extract-mode independently from each other. A total of 22 bond-pads were used to control the extract states (11 bond-pads for each group of 8 channels; 4 bond-pads for channel #1 and 7 bond-pads for channels #2 to #8). Although multiple channels could also be selected to extract data, due to the channel priority, the channel closest to the electrical output port was the one that accessed the port.

To simplify the chip as much as possible, the address comparison circuitry was omitted from the design. This was also in part due to a lack of extra electrical output bond-pads. However, the segmentation of the first channel in each of the 2 groups of 8 channels would allow for a similar, yet slower, form of address comparison. Since the 4 chips in the optical system could simultaneously transmit on different segments of the first channel, the activity of the 4 chips in the system could be continually monitored. This part of the chip operation was suggested by Prof. Ted Szymanski but was not fully explored at the time of writing this thesis. Therefore, the details of this communication method were not fully determined. The P3B chip was normally tested with prior knowledge of the source and destination of the data, and thus the control lines were set before the data transmission.

The P3B chip did not require a serial-to-parallel register, the only two control bits were for the bit-inversion multiplexers and the synchronous/asynchronous behaviour of the smart pixels. Both these control signals were supplied using 4 bond-pads (a two “bit-inversion” bond-pads and two “asynchronous” bond-pad for both the upper-group and the lower-group). Finally, since the smart pixels could be placed into asynchronous mode and because the design was fairly simple, “test-bits” were not required on the chip.

The P3B chip had an additional set of structures that were used to obtain direct measurements of the MQW diodes. These structures could also possibly provide a method to drive the optical links faster than the 1.5- μm CMOS technology would permit. For testing purposes, 4 clusters (2 modulator clusters and 2 detector clusters) were set apart from the smart pixel array design. Although these clusters reduced the number of smart pixels in 4 channels (from 16 smart pixels down to 8 smart pixels), direct access to the MQW diodes for testing purposes was essential. These clusters were known as the “hardwire” clusters. The two “hardwire” modulator clusters were composed of 8 dual-rail pairs of MQW diodes where the “center-tap” point was routed directly to passive bond-pads [Figure 4-24]. There were no transistors between the dual-rail totem poles and the bond-pads and thus the MQWs could be driven at whatever speed could be supplied to the chip. Proper line termination of these bond-pads had to be designed and is part of the P3B packaging assembly designed by Mr. Michael Ayliffe and Mr. Alan Chuah [ref 11]. The “hardwire” detector clusters were composed partly of single-ended MQW diodes that were connected directly to passive bond-pads and subsequently attached to off-chip transimpedance amplifiers. The other diodes in the “hardwire” detector clusters were dual-rail totem pole MQW diodes with on-chip transimpedance amplifier optical receivers that were directly routed to passive bond-pads. Using these structures, initial MQW diode tests could be made as well as in-system tests of the speed, latency and bit-error rates of the optical link.

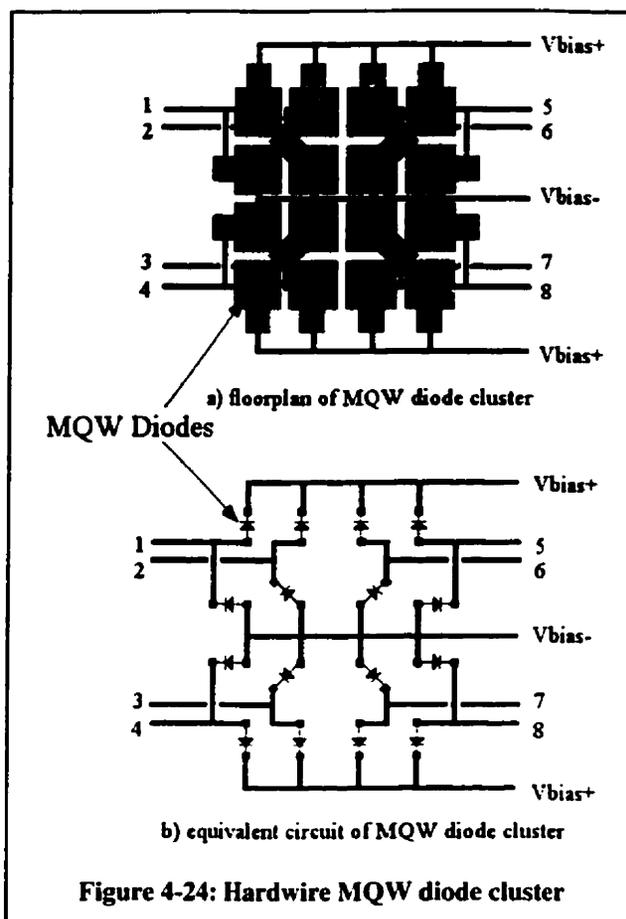


Figure 4-24: Hardwire MQW diode cluster

There were a total of 232 bond-pads on the P3B chip. There were 24 ground pads, and 20 power pads. There were 16 bond-pads for MQW diode biasing and 2 bond-pads for transimpedance amplifier biasing. A total of 32 bond-pads were used for the “hardwire” clusters, and a total of 24 were used for active alignment techniques. The remaining 114 pads were for the digital access to the smart pixel array.

4.5.3) Layout

The layout of the P3B chip was very similar to the layout of the P3A chip. Columns of clusters had to alternate between modulating clusters and detecting clusters on the same pitch as the P3A chip. Therefore, the smallest repeatable region was the “super-cluster” with an area of 1600- μm x 800- μm containing both types of cluster. It contained 16 modulator devices and 16 detector devices for a total of 8 smart pixels (assuming dual-rail MQW diode structures). With 4 horizontally placed “super-clusters”, 2 horizontal channels of 16-bits were defined.

Since the technology was limited to 2 metal layers, the routing of power, ground, MQW bias, control signals and data signals was slightly more complicated. Area of the super-cluster was required for power and bias line routing, whereas in the P3A chip these lines could rest on top of each other in different metal layers. Therefore, each smart pixel had to be 80- μm high and 450- μm wide and stacked on top of each other allowing the top and bottom of the super-cluster to route power and MQW bias lines horizontally along the chip. All smart pixel transistors were sized with a balancing width ratio ($W_P = 2.5W_N$ for the 1.5- μm CMOS process) and a minimum transistor width of 2- μm to allow equal rise and fall times. A much more strict approach of using metal-1 lines for all horizontal interconnects and metal-2 lines for all vertical interconnects was applied. Although, the same basic methodology from the P3A chip was still used.

Since there were only 2 metal layers, no circuitry could be placed below the clusters. The cluster biasing lines and the attachment points required metal-2 (or top-metal) making it very difficult to place circuitry below the cluster.

In the P3B chip design, the clock distribution was very rigorously planned. Although the P3A chip had a similar clock structure, the P3B was truly an equal length,

binary tree. All paths from the clock bond-pad to each of the smart pixels were the same length and passed through the same number of buffer elements.

4.6) Optoelectronics

The MQW diode was the enabling optoelectronic technology for all four chips. A reflection-mode MQW diode can modulate an incoming constant optical beam by changing its absorption as a function of applied voltage. The reflected light is encoded with the electrical data as high and low intensity pulses. A major benefit of the MQW diode was that it could be used as a detecting device as well and thus both modulating and detecting devices could be processed in one step. A major disadvantage of the MQW device was that it required an external, constant light source to “read-out” the state of the device. A second drawback was due to the type of optical encoding used. For the chosen encoding scheme, twice the number of optical beams were required. Therefore, two MQW diodes were used to encode each “optical” bit from the smart pixel. The reason for this redundancy was due to the poor contrast ratio of the MQW device. The MQW did not have a true “off” state, it only changed in reflectivity from a “high” intensity to a “low” intensity making it difficult to construct a *simple* receiver circuit that was based on a single beam of light. The method chosen was called dual-rail encoding and is similar to the way differential emitter-coupled logic (ECL) uses both the true value and its complement to overcome common-mode noise in its data transmission.

This chapter will begin by discussing the structure and operation of the MQW, followed by the structure of a typical MQW diode. A brief discussion of the integration of the GaAs MQW diodes with the silicon chip will follow. Typical curves for the absorption of MQW devices will be presented, showing their dependence on both voltage and temperature. Next, two types of detector circuitry will be presented, as well as a typical driver circuitry for the modulator MQW devices.

4.6.1) The multiple-quantum-well

The multiple-quantum-well (MQW) diode exhibiting the quantum-confined-STARK-effect (QCSE) was used as the optoelectronic device in the previously described

VLSI-OE chips [ref 12]. The MQW is based on gallium arsenide (GaAs), which is a direct band-gap semiconductor. The position of the band extrema of a direct bandgap semiconductor, such as GaAs, allow for the creation of a pseudo-particle called an exciton [figure 4-25].

Since the valence band maximum and the conduction band minimum are located “near” each other (in terms of similar momentum vectors), a condition where a single electron in the conduction band can be attracted to a single hole in the valence band becomes possible [figure 4-26]. This particle interaction is called an exciton pair, and exists for only brief moments in a bulk GaAs semiconductor. An exciton is similar to a hydrogen atom where the electron revolves around the nucleus. In an exciton, the electron and hole revolve around each other up to about a 300-Å radius.

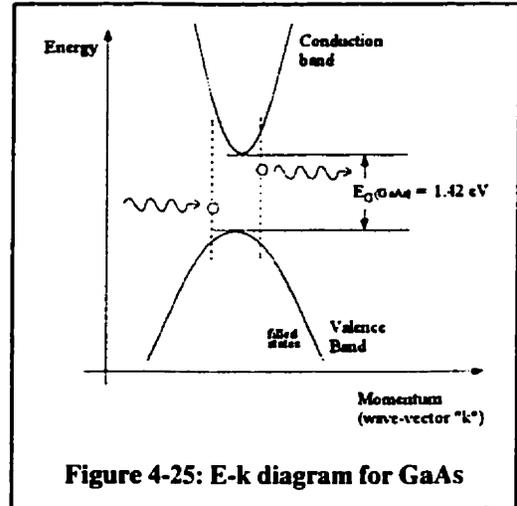


Figure 4-25: E-k diagram for GaAs

In a bulk GaAs semiconductor, the exciton behavior is difficult to observe (except at very low temperatures). The absorption curve of bulk GaAs as a function of wavelength (at room temperature) is shown here [figure 4-27], where the x-intercept is approximately the energy of the band-gap. Gallium arsenide requires a minimum energy of 1.422-eV at 300K (an equivalent wavelength of 872.89-nm) to raise an electron into the valence band to the conduction band. Obviously, any energy greater (or wavelength shorter) than this will raise the electron as well. The exciton can only exist for brief moments in a bulk crystal because the

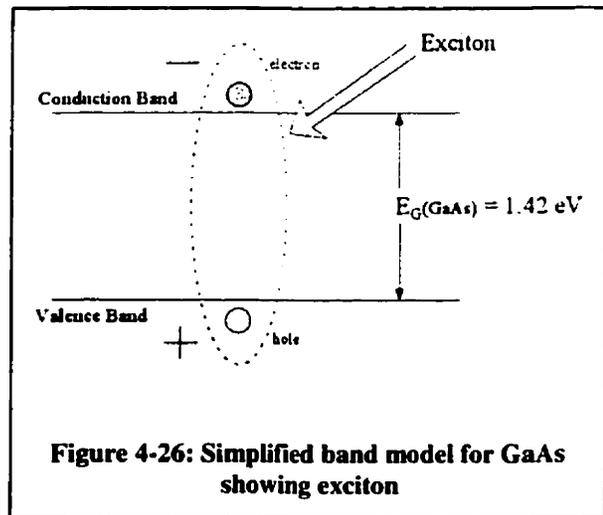
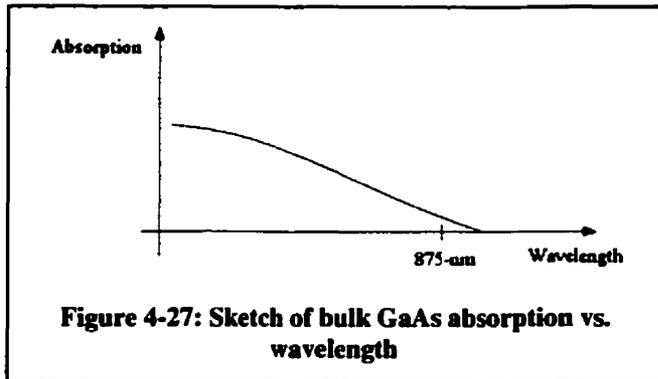


Figure 4-26: Simplified band model for GaAs showing exciton

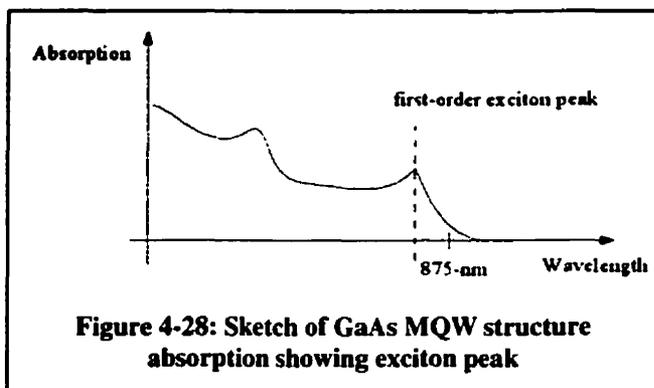
electron-hole electronic attraction is very weak, and other particle interactions and lattice vibrations disrupt the attraction so it is not apparent in the absorption spectra. In order to

increase the likelihood of an exciton existing, the quantum-well is used to confine the particles in place.



The quantum-well is a well-studied phenomenon and can be modeled to a first approximation by the “particle-in-a-box” theoretical experiment. By layering GaAs and Aluminum-GaAs (AlGaAs) in very thin layers roughly 50-Å thick, potential barriers for the electron

and hole are created similar to the “particle-in-a-box”. These barriers do not allow the electron or hole to physically move away from each other; they become confined to a



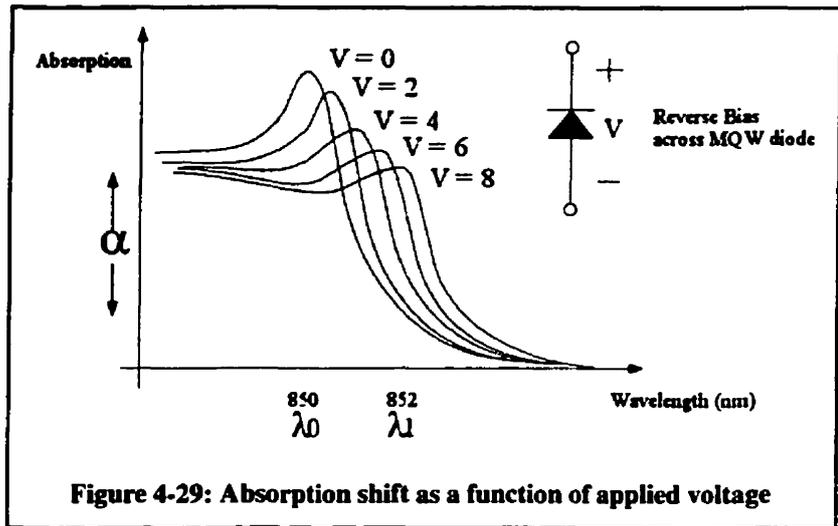
region of space that allows the exciton to exist for a longer duration and in more abundance. When multiple layers of GaAs-AlGaAs are used, the effect becomes even stronger and the absorption spectrum begins to change shape with a very predominant absorption

peak near the GaAs band-gap energy [figure 4-28]. Other exciton orders are also created at higher energies just as higher modes are predicted in the “particle-in-a-box” experiment, but these are of no importance in this discussion. However, it is important to realize that the exciton is only a pseudo-particle and does not exist until some mechanism (such as high energy light) creates the electron-hole pair. The fabrication of the multiple quantum-wells simply increases the probability that the exciton will exist, but does not actually create them.

The GaAs-AlGaAs multiple quantum-well is the structure that enhances the exciton peak, but the key to light modulation is the ability to dynamically move this peak in wavelength as a function of applied external voltage. For example, when a narrow line-width laser with the same wavelength as the exciton peak is directed at the MQW stack, a

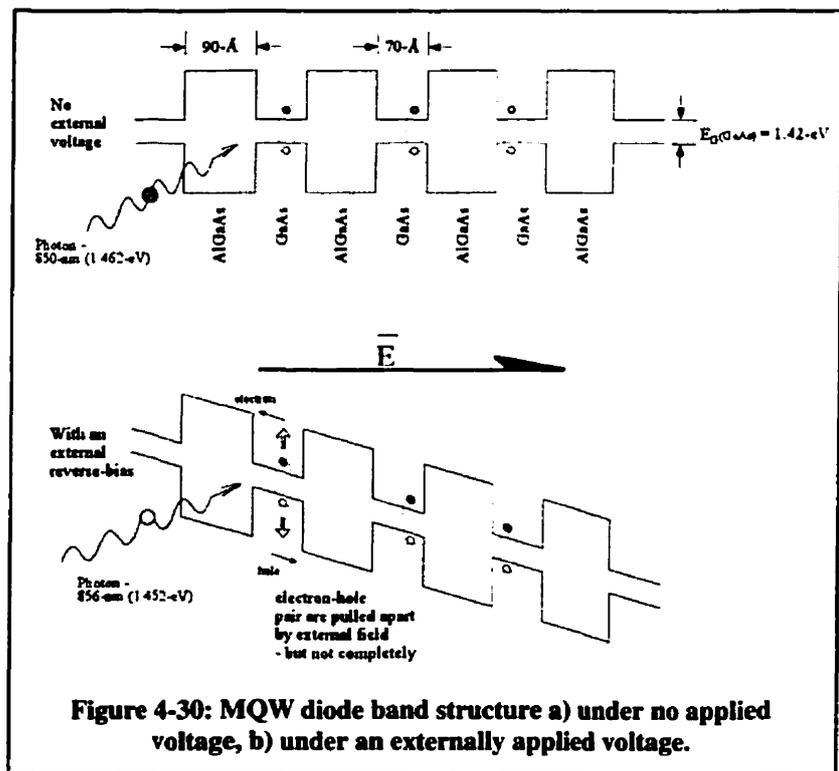
substantial amount of light energy can be absorbed. If the exciton peak is now shifted to a shorter wavelength (by means of an externally applied voltage), the incident light would not be absorbed as much [figure 4-29]. This effect is known as the quantum-confined-STARK-effect (QCSE) and relates the device absorption to applied electric field.

By fabricating a doped n-type GaAs region on one side of the MQW-stack, and a p-type GaAs region on the other side of the MQW-stack, a P-i-N diode is formed where the “intrinsic” region is composed of the GaAs-AlGaAs layers. When the diode is reverse-biased, the constant electric field



across the MQW-stack changes the energy required to form the exciton. Since the AlGaAs barriers continue to confine the exciton spatially, the external electric field cannot pull the electron-hole pair apart completely [figure 4-30].

However, with the added energy of the external field, a less energetic photon is required to create the exciton because the external field is helping to create the electron-hole pair. This results in a “red-shift”



in the absorption spectra allowing longer wavelength (or lower energy) light to create the exciton peak. The exciton “red-shift” is similar to the shift seen in hydrogen atom absorption spectra under strong external electric fields.

4.6.2) The MQW device patterning

Although all four VLSI-OE microchips used the MQW modulator device, the structure and geometry of the devices differed between Phase-II and Phase-III. The first two chips, the Beta-Chip and the Workshop chip, used MQW structures proprietary to Lucent Bell Labs (now Lucent Technologies). The Phase-III-A chip used a structure designed in part by Prof. F.A.P. Tooley at McGill University (from Heriot-Watt University in Edinburgh, Scotland) and Dr. Anthony Springthorpe of Nortel Semiconductors. The mesa patterning and virtually all components of the fabrication for the Phase-III design were under the direct control of the design team at McGill. The design used in the Phase-III-B chip was very similar to the Phase-III-A design, except that the structure was grown and designed by Dr. John Trezza.

The structure of the Phase-III-A MQW growth is given in [figure 4-31], and is representative of any of the other structures used. The structure was grown up side down with respect to the final orientation of the device once attached to the silicon CMOS processing chip.

The layers were grown on a GaAs substrate and essentially consisted of the following regions. The first region was the bulk GaAs substrate, followed by a buffer layer, and an etch-stop layer. A Silicon (Si) doped GaAs and AlGaAs layer made up the n-type region, then the quantum-well stack was grown and finally the Beryllium (Be) doped GaAs made up the p-type region. The quantum-well layers were deposited using molecular-beam epitaxy, which grew 60 alternating layers of GaAs (90-Å thick) and AlGaAs (35-Å thick) for a total multiple quantum-well thickness of 0.75- μm .

Once the layers were grown on the wafer, the appropriate patterns were etched into the GaAs layers. In the first two chips, the processing and the geometry were specified by Lucent. On the second two chips, both the processing and the geometry were under the control of the McGill Photonic Systems group. Although the exact processing

procedure of the Lucent grown devices was proprietary, the overall concept used to grow and pattern the devices was similar to the method used to construct the Phase-III-A devices.

The method used to grow and pattern the Phase-III-B devices was also similar to Phase-III-A, except a dry-etch technique was used. The following will briefly describe the wet-etch process developed by Dr. Edwis Richard at L'Ecole Polytechnique in Montreal Canada (currently in France) used in the Phase-III-A design, where the end result was to create a two-contact P-i(MQW)-N diode.

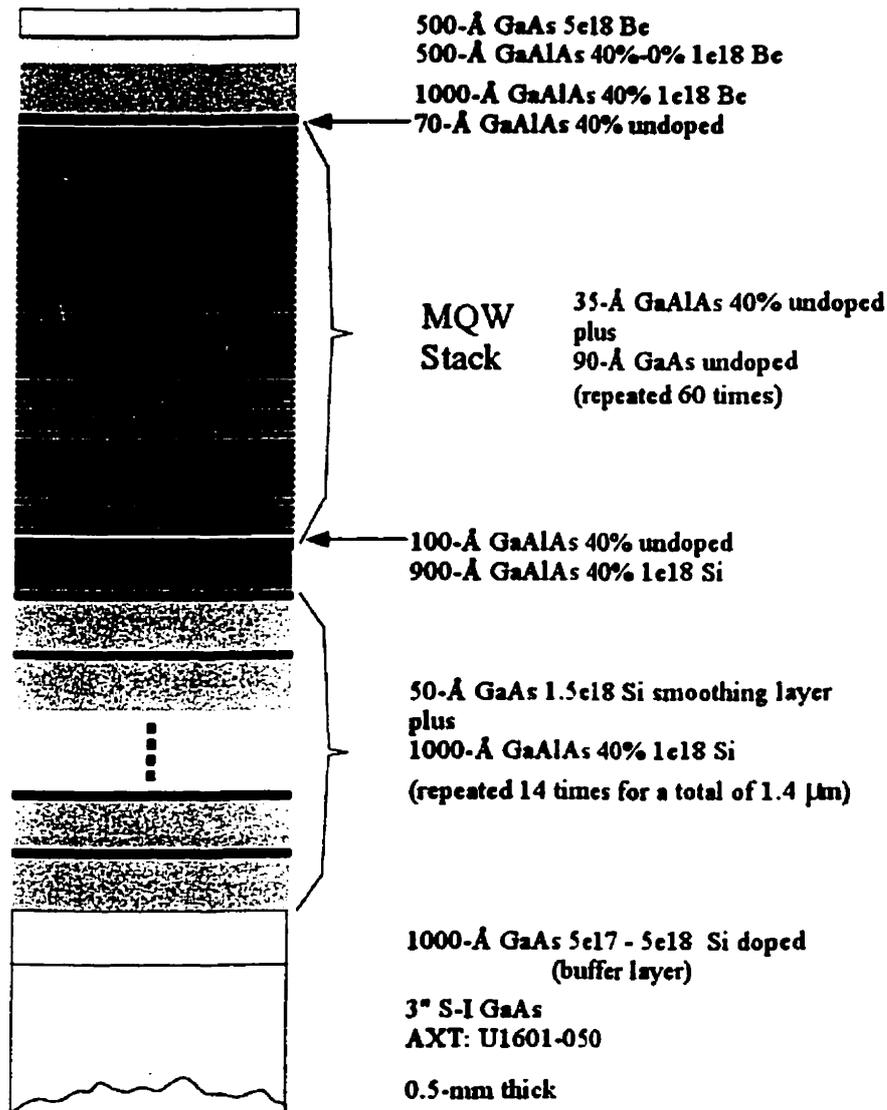
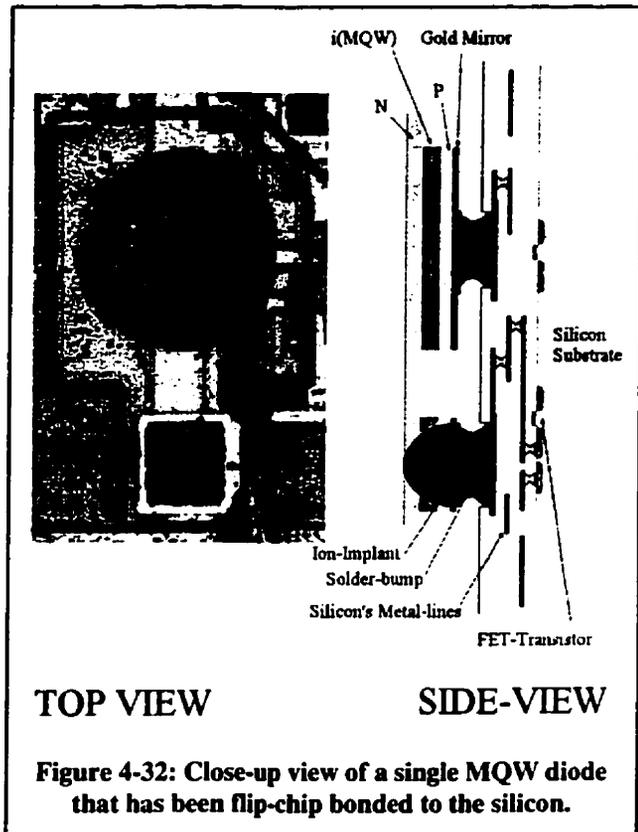


Figure 4-31: Sketch of the device structure used for the P3A (and P3B) MQW diodes.

The first step in constructing the diode was to create two closely spaced mesa structures (or islands) of the grown layers. This pair of mesa islands would ultimately become the n and p contacts for each diode as well as provide mechanical support of the diode once flip-chipped onto the silicon chip [figure 4-32].

The second step was to circumvent the p-type region of one of the mesa islands so that a direct connection to the n-type region could be made. This was done by selectively coating one of the mesa islands with AuGe (Gold-Germanium alloy). In the other processes, this was done using ion-implantation through the entire stack until it reached the n-type region. The p-type region was then coated with a Ti/Pt/Au (Titanium, Platinum, and Gold layers), this formed the metal contact as well as the "internal" mirror surface at the Gold-GaAs interface [figure 4-33]. Each mesa-pair was then

isolated from other mesa pairs by etching all the way down to the etch-stop layer. A diffusion barrier was then formed over both contacts to prepare then for the solder



TOP VIEW

SIDE-VIEW

Figure 4-32: Close-up view of a single MQW diode that has been flip-chip bonded to the silicon.

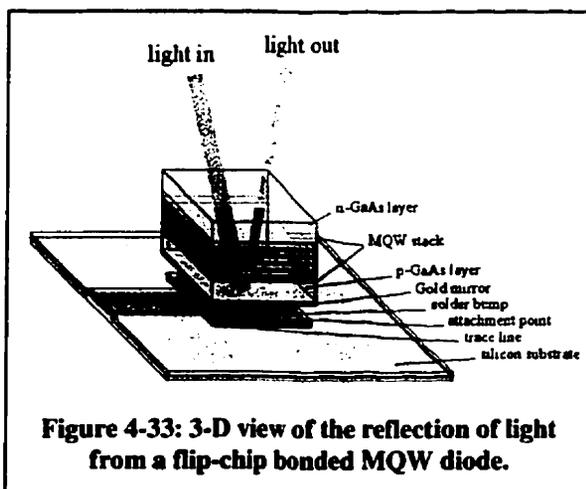


Figure 4-33: 3-D view of the reflection of light from a flip-chip bonded MQW diode.

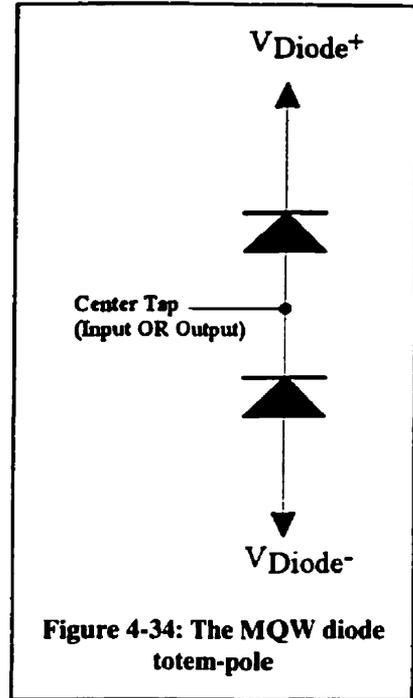
deposition and prevent the solder from diffusing into the diode. The final step was to pattern each contact point with solder and flip the GaAs substrate with its etched diodes up side down. The GaAs chip (or wafer depending on the technique) was then aligned and placed in contact with the corresponding contact points on the silicon CMOS processing

chip. A fluid epoxy was injected into the space between the two chips to add to the mechanical support, and the whole assembly was then heated and cooled so that the solder could wick to both the silicon and GaAs contact points. The entire GaAs substrate, which was about 500- μm thick (about 98% of the total thickness of the layer-grown wafer) was then lapped down until the etch-stop layer. This step completely isolated the diodes from one-another. The top-surface of the device (with respect to the silicon CMOS chip) was the n-type region. This surface was A/R-coated (single layer, narrow-band, anti-reflection coating) for the 852-nm wavelength used in the system.

4.6.3) MQW device operation

As discussed earlier, the MQW diode is used as a modulation device. The optical design required that the constant power optical beam was normally-incident on the active region (quantum-well region) of the device to “read-out” the state of the absorption (either high absorption or low absorption as a function of applied voltage). The MQW device was used in reflection-mode. Light traveled through the quantum-well stack and hit the gold backing-mirror. The light then reflected back through the quantum-well stack towards the source. If the stack was in a state of high absorption, very little light would be reflected back. The expected absolute reflectivity in each of the extreme states was 30-% in the high absorption state and 80-% in the low absorption state. Using these two extremes in optical power reflection, an optical bit pattern could be produced by modulating the voltage. To operate the MQW, the voltage bias is changed such that the exciton peak is moved “into” and “out-of” the operating laser light wavelength (usually 852-nm). Due to the nature of the exciton, a change in absorption at 852-nm can be obtained in two ways. The first method uses the λ_0 wavelength and takes advantage of the peak from the left side [figure 4-29]. The second method uses the λ_1 wavelength and takes advantage of the much longer slope to the right of the peak. The second method offers much better contrast than the first and was used in all of the systems designed.

The fact that the device did not have a true “off” state complicated the type of optical transmitter and optical receiver circuit on the chip. Furthermore, the precise reflectivity of each device could be slightly different from one-another, which complicated the optical transmission even further. The low contrast ratio and the variation between devices were overcome using a technique used in earlier photonic switching demonstrator systems at Lucent [ref 13]. This approach was a dual-rail totem-pole approach, where each optical bit was encoded using two MQW devices in complementary states. This circuit originated through the work with symmetric-SEEDs (self electro-optic effect devices) and all optical switching techniques [ref 14]. The totem-pole diodes were reverse biased by applying a voltage across the two series connected



diodes. The bias voltages were nominally: $V_{Diode+} = +8.3\text{-V}$ and $V_{Diode-} = -3.3\text{-V}$ [figure 4-34]. For optical data transmission, a signal line (0-V or 5-V) was attached to the middle of the series connected diodes. The voltage across the top device would be the complement of the bottom device and therefore one device would be in a high absorption state while the other would be in a low absorption state. If a pair of MQW devices had a high/low absorption pair, a logical-high bit was interpreted. When the MQW devices had a low/high absorption pair, a logical-low bit was interpreted. This technique is analogous to the complementary signaling used in some ECL (emitter-coupled logic) circuits [ref 15]. By using pairs of adjacent MQW devices, redundant optical paths were required. However, because the bits were encoded using the “sign” of the complemented optical signals, very simple (and fast) on-chip amplifiers could be constructed that amplified the difference in the pairs of optical beams. The main advantage with dual-rail signaling was that it did not require a reference voltage to compare against. A single beam of light (especially one that does not completely turn off) would require a corresponding voltage to compare against to make a decision as to the value of the bit (either 1 or 0). In an

optical system, where the reference power may vary from chip-to-chip, the dual-rail system is optimal. The suppression of common-mode noise is also an advantageous characteristic of the dual-rail approach because the common attenuation of a pair of beams can be cancelled-out by an appropriate receiver amplifier.

An additional concern of the MQW device is its sensitivity to temperature changes. Just as an increase in external electric field reduced the minimum amount of energy required to create an exciton pair, so does an increase in thermal energy. The exciton peak can be “red-shifted” using temperature, allowing lower energy (longer wavelength) photons to create excitons. The exciton peak moves approximately 3-nm/°C [ref 16]. Therefore, rather stringent thermal management techniques were used to remove the heat from the surface area of the chip [ref 17]. The estimated operating temperature of the silicon CMOS circuitry was used as a factor when designing the quantum-well thickness, so that the exciton peak would lie in the correct part of the absorption spectrum.

4.6.4) MQW device model

One of the key factors in proving the viability of this technology is the exceedingly high off-chip (or off-board) data rates attainable using micron-sized optoelectronic devices. Therefore, the device size and related internal circuit model were critical. The load that the optoelectronic device offered to the driving circuitry would limit its overall bandwidth.

To a first order approximation, a reverse-biased P-i-N diode can be modeled as a parallel-plate capacitor with the p-type region and the n-type region separated by the intrinsic MQW region. For the Phase-III-A MQW diodes, the area was approximately $3848\text{-}\mu\text{m}^2$ (a diameter of $70\text{-}\mu\text{m}$), with a plate separation of $0.8\text{-}\mu\text{m}$. The dielectric constant of the GaAs-AlGaAs layer was approximately; $K_{\text{GaAs}} = 3.5$. Therefore, the capacitance would be roughly 150-fF. The circuit model also can be expanded to include other effects, such as parasitic capacitance and contact resistance. One model developed for the MQW device in reverse bias is a detailed analysis of each region of the device [ref

18] and predicts $0.11\text{-fF}/\mu\text{m}^2$. This model predicts as much as 500-fF of capacitance for each $70\text{-}\mu\text{m}$ diameter MQW device [figure 4-35].

Another part of the model was the mechanism used to introduce optically generated photocurrents. When the model was used to simulate transmitter performance, a pair of constant current sources were used in each diode model to simulate the photocurrent generated by the “read-out” optical beams on the diodes. The optical power expected on the modulators was roughly $100\text{-}\mu\text{W}$ which could generate a photocurrent of $50\text{-}\mu\text{A}$. When the model was used to simulate the performance of the receiver circuitry, the constant current

sources were replaced with modulated current sources with a minimum and maximum of typically between $10\text{-}\mu\text{A}$ and $5\text{-}\mu\text{A}$.

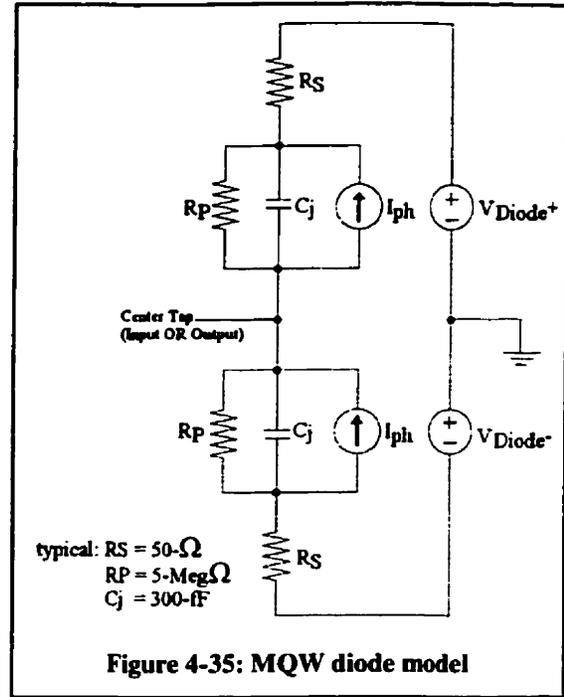


Figure 4-35: MQW diode model

4.6.5) MQW transmitter circuit

The transmitter driver circuits in each of the four chips were based on scaled CMOS inverters capable of driving large output capacitances. The method is identical to the scaling procedure outlined in [ref 19] used to match output drivers with output bond-

pad capacitances. Given that the propagation delay through a minimum sized inverter is; t_d . And that the propagation delay through an inverter

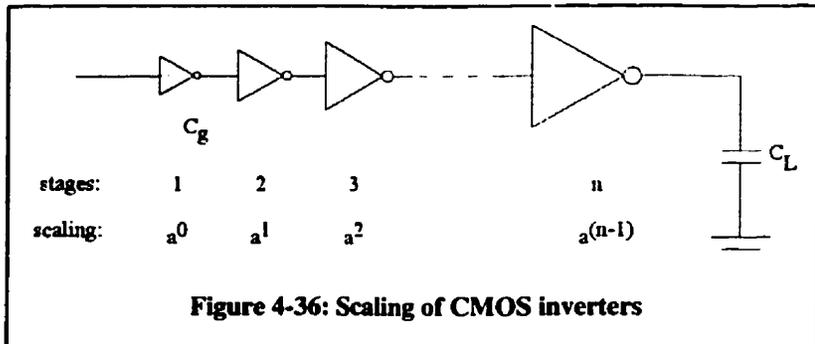


Figure 4-36: Scaling of CMOS inverters

which is 'a' times larger than the minimum would be; at_d (assuming the minimum size

inverter is driving the larger inverter). Furthermore, if there are 'n' inverters connected in series [figure 4-36], where each inverter is 'a' times larger than the previous one, the total delay is given by; nat_d . Next, assuming that the series of inverters had a final load capacitance, C_L , and that the input capacitance of the first minimum sized inverter was; C_g . The ratio of load capacitance to inverter capacitance would be; $R = C_L/C_g$. This ratio must also be the same as the area scaling ratio; $R = a^n$. The total delay is then given by [Eqn. 4-1]:

$$\begin{aligned}
 \text{Eqn. 4-1:} \quad \text{Delay}_{\text{TOT}} &= nat_d \\
 &= [\ln(R)/\ln(a)]at_d \\
 &= t_d \ln(C_L/C_g) a/\ln(a)
 \end{aligned}$$

The total delay is minimized when $a = 2.71$ (or $a = e$). The number of stages is easily calculated using; $n = \ln(R)$, where 'n' is rounded to the nearest integer.

The first step in the transmitter design was to match the nMOS and the pMOS transistors to produce a symmetric inverter transfer function. Using HSpice, an appropriate transistor model was used (in this case, the model was a 5-V, level-3 empirical model obtained from the 1.5-micron MITEL fabrication run through CMC). The nMOS transistor was simulated with a gate-length of $L_N = 1.5\text{-}\mu\text{m}$ and gate-width of $W_N = 2.0\text{-}\mu\text{m}$, and the pMOS transistor was simulated with a gate-length of $L_P = 1.5\text{-}\mu\text{m}$ and gate-widths from $W_P = 2.0\text{-}\mu\text{m}$ up to $W_P = 10.0\text{-}\mu\text{m}$. A symmetric transfer function was produced when the pMOS transistor had a gate-width of $W_P = 5.0\text{-}\mu\text{m}$. Using the correct lengths and widths, the rise-time and the fall-time of the inverter was obtained: $t_r = 0.189\text{-ns}$ and $t_f = 0.151\text{-ns}$. Also, the average propagation delay-time through a minimum sized inverter was $t_d = 0.128\text{-ns}$.

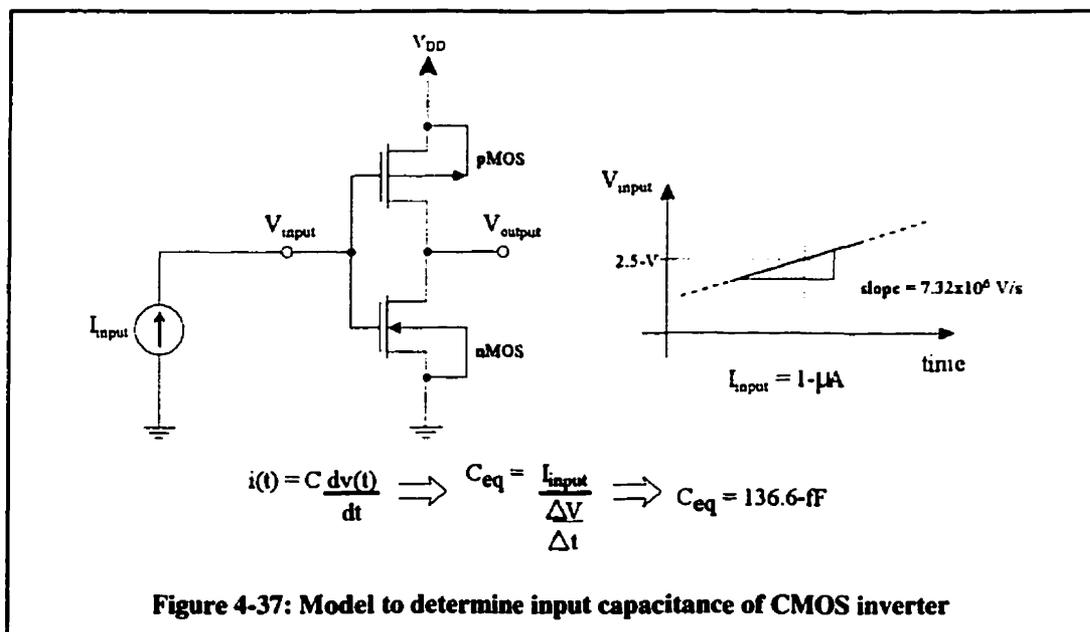
The second step in the design was to calculate the total input capacitance of a minimum sized CMOS inverter. The total gate capacitance is given by [Eqn. 4-2]:

$$\text{Eqn. 4-2:} \quad C_{\text{cq}} = (C_G)_N + (C_G)_P + (1 - A)((C_{\text{DG}})_N + (C_{\text{GD}})_P)$$

Where $(C_G)_N$ and $(C_G)_P$ can be obtained adding the gate-to-source, C_{GS} , capacitance to the gate-to-body, C_{GB} , capacitance of the respective transistors. The gate-to-body capacitance can be estimated using the equation for the parallel-plate capacitor on the polysilicon-oxide-semiconductor gate structure [Eqn. 4-3]. This value is then added to the parasitic gate-to-source capacitance, C_{GS} , which can be found in the output results from the HSpice simulation (for this analysis, $(C_{GS})_N = 1.876$ -fF and $(C_{SG})_P = 4.916$ -fF). This produced a gate capacitance of $(C_G)_N = 5.4544$ -fF and $(C_G)_P = 13.8621$ -fF.

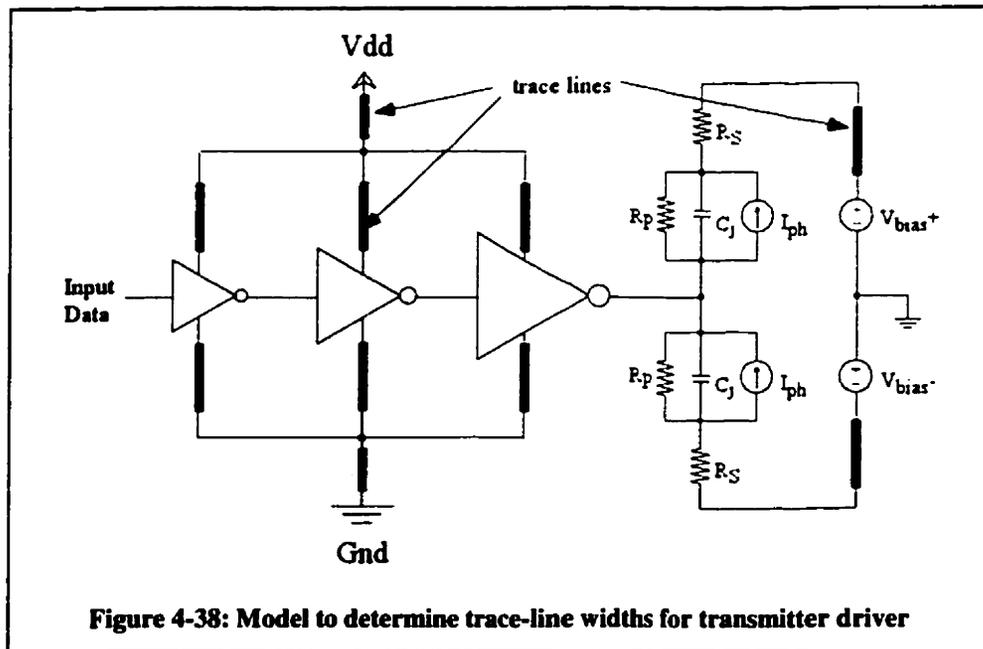
Eqn. 4-3: $C_{GB} = K_{SiO2} \epsilon_0 \text{Area} / t_{ox}$

The gate-to-drain capacitances for both transistors can be added but must also be modified by the open-loop gain 'A' due to the Miller-effect feedback. For this example, $A = -25.77$ -V/V, $(C_{DG})_N = 1.3233$ -fF, and $(C_{GD})_P = 3.1123$ -fF. The areas of each gate were $\text{Area}_N = 3.0$ - μm^2 and $\text{Area}_P = 7.5$ - μm^2 . Therefore, the total capacitance is $C_{eq} = 134.6$ -fF.

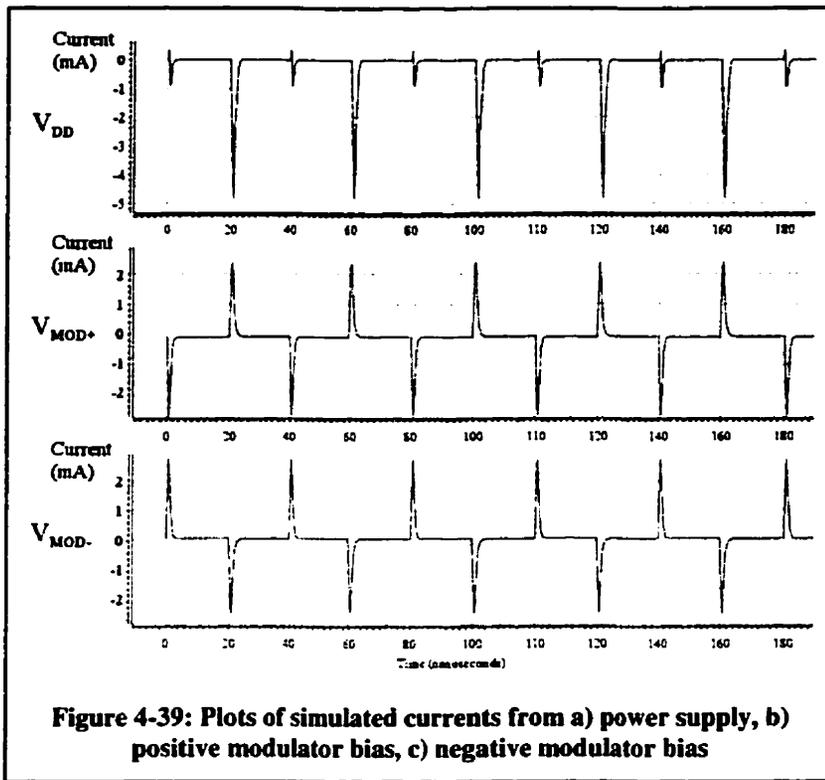


If the circuit in [figure 4-37] is used, a more precise input capacitance can be determined. When the input node is initialized to its 'dc' operating point ($V_{IN} = V_{OUT} = 2.530298\text{-V}$), the input current can be divided by the slope of the voltage vs. time of the input node to calculate the input capacitance; $C_{eq} = I_{IN}/\text{slope} = 136.6\text{-fF}$. Both methods produce results in good agreement.

The values for the minimum inverter input capacitance, the delay-time, the transistor sizes, as well as the estimated MQW capacitive loads can now be used to determine the parameters of the scaled inverter chain. With a value of $C_L = 1000\text{-fF}$, a value of $C_{eq} = 135\text{-fF}$, and a delay time; $t_d = 0.128\text{-ns}$. The minimum number of stages required is $n = 2$ with $a = 2.71$. However, the best performance was obtained with $a = 3$ and $n = 3$, although this did consume more dynamic power. In a technology that had a smaller gate-width (such as the 0.8-micron technology of the Phase-III-A chip), the minimum-sized inverter capacitance was significantly smaller and therefore required a larger number of staged inverters to drive the same MQW capacitance. The number of stages could be as much as $n = 3$ or $n = 4$ with a larger scaling factor of $a = 3.2$ for smaller technologies.



An HSpice simulation of the transmitter circuit shown in [figure 4-38] allowed an estimate for the minimum number of bond pads for both power and ground as well as for the MQW bias lines. A square-wave input pulse-train with a frequency of 25-MHz was applied to the input stage of the transmitter circuit. When the currents of the bias lines



were plotted, the average and peak currents were obtained: $I_{\text{PEAK-MOD}} = 2.5\text{-mA}$ and $I_{\text{AVE-MOD}} = 0.10\text{-mA}$ [figure 4-39]. These currents, when averaged, cancel-out due to the symmetric current pulses above and below the static current of $50\text{-}\mu\text{A}$ due to the incident light. To obtain an average

current with which to calculate the trace line widths, the current pulses were “rectified” and time-averaged. This resulted in a time average current of 0.19-mA for both the positive and negative bias lines. Multiplying by the total number of devices in the array (256 smart pixels in the Phase-III chips), the total time average current were: $I_{\text{AVE-MOD-TOT}} = 48\text{-mA}$ and $I_{\text{PEAK-MOD-TOT}} = 640\text{-mA}$ for each bias line. Allowing for a factor of 2 safety margin, 4 bias line bond pads were required for each bias line since a single bond pad could only handle 30-mA average current. The current pulled from the power supply, V_{DD} , was also plotted and the average and peak currents were calculated. The time average current from V_{DD} was: $I_{\text{AVE-VDD}} = 0.425\text{-mA}$ and $I_{\text{PEAK-VDD}} = 5\text{-mA}$. Again, multiplying by the total number of devices in the array, the total currents were calculated: $I_{\text{AVE-VDD-TOT}} = 108.8\text{-mA}$ and $I_{\text{PEAK-VDD-TOT}} = 1280\text{-mA}$. Again, allowing for a factor of 2 safety margin, at least 8 V_{DD} bond pads (and 8 Ground Pads) were required just for the

power consumed by the transmitter. Since space permitted, additional bond-pads were provided for both power and bias. Although the receiver circuit required much less current from the MQW diodes, the same number of bond pads for detector MQW biasing was used.

4.6.6) MQW detector circuit

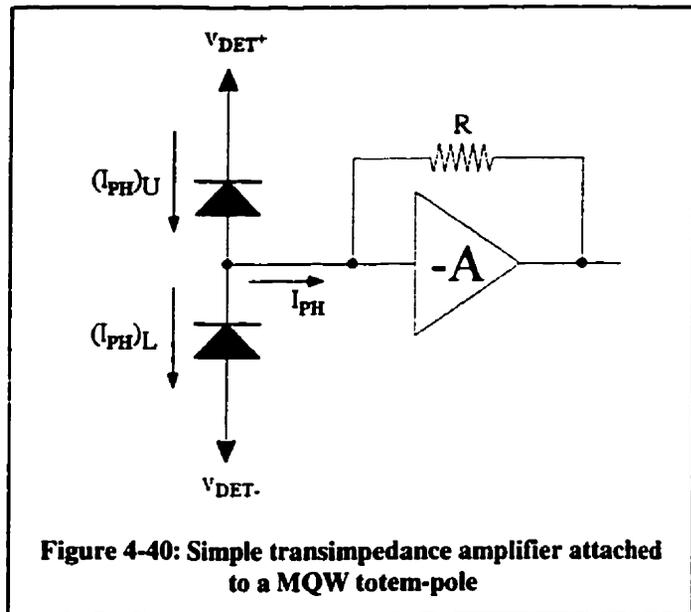
The most beneficial characteristic of the MQW P-i-N diode is that it can function as both a modulator of light and a detector of light depending on its voltage-bias and operating conditions. If the reverse-bias voltage across the P-i-N diode is kept constant, the device can generate a photocurrent directly proportional to the incident light power. The voltage across the terminals varies only slightly as a function of photocurrent. As discussed in a previous subsection, the dual-rail nature of the optical data required that the optical receiver circuit was capable of using the redundant information effectively. In the following subsection, two types of optoelectronic receiver circuits will be discussed. The transimpedance optoelectronic receiver circuit was used in the Beta-Chip, Workshop-Chip, and the Phase-III-B chip. The charge-sense amplifier (CSA) was used in the design of the Phase-III-A chip. Not only did the amplifiers have to achieve high-bandwidth and high-gain, but the circuit had to be replicated up to 256 times for all the smart pixels on the chip without taking any significant amount of chip area.

4.6.6.a) The TZA amplifier:

The transimpedance (TZA) amplifier is ideal for the detection of small photocurrents at high data rates and has been the receiver design used in numerous VLSI-OE designs [ref 20]. A TZA can maintain a high bandwidth because it does not significantly load the input stage with a high resistance, allowing the RC-time constant to remain low. The TZA amplifies small changes in input current and matches the input stage to larger input impedance amplifier stages. The small optoelectronic photocurrents from the MQW devices are used to generate 0-to-5 Volt digital logic swings. Assuming that an input photocurrent of 5- μ A corresponds to a digital output voltage of 5-V, a gain of 1,000,000-V/A is required. However, because a single-stage open-loop amplifier with 1,000,000-V/A of gain can be unstable, the amplification is performed in several stages,

where the closed-loop feedback of the TZA is used at the most critical stage; the interface to the MQW photodiodes. The pairs of optical beams are detected by two MQW diodes in the same type of dual-rail totem-pole configuration as the transmitter and the center-tap of the totem-pole is used as the input to the TZA [figure 4-40]. The photocurrent generated in the upper MQW diode, $(I_{PH})_U$, and the photocurrent generated in the lower MQW diode, $(I_{PH})_L$, must obey Kirchoff's current law at the input node. Therefore, the current into the TZA is given by: $I_{PH} = (I_{PH})_U - (I_{PH})_L$. This configuration allows the TZA to amplify only the difference in the complementary optical signals. It is also possible to reject common-mode optical noise (such as power fluctuations) since the signals are subtracted from one another.

The most common TZA is the inverting operational-amplifier (OP-AMP) [ref 21]. The circuit can be generalized even further by including a dc-level shift at the positive-terminal and a finite gain 'A' [figure 4-41]. The following equations are used to relate the input photocurrent to the output voltage [Eqn. 4-4, 4-5, 4-6]:



Eqn. 4-4: $V_{DC} + V_{OP} = I_{PH}R_F + V_{OUT}$

Eqn. 4-5: $V_{OUT} = -AV_{OP}$

This produces the relationship:

Eqn. 4-6: $-(1+1/A)V_{OUT} = I_{PH}R_F - V_{DC}$

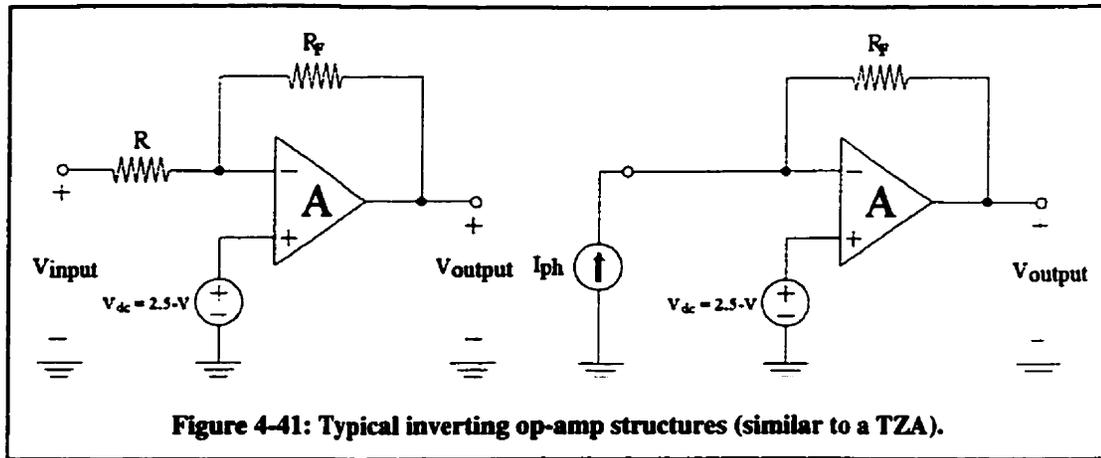


Figure 4-41: Typical inverting op-amp structures (similar to a TZA).

In equilibrium, when $I_{PH} = 0-A$, V_{OP} must be zero (when 'A' is infinitely large). Also, the voltage across the feedback resistor must be zero and the output must be V_{DC} . A plot of the transfer function is given here [figure 4-42] where the current-to-voltage gain is essentially the value of the feedback resistor; $V_{OUT}/I_{PH} = -R_F$ (plus a dc-shifted value).

The CMOS inverter-amplifier with a pMOS feedback resistor [figure 4-43] is very similar to the inverting OP-AMP configuration. The transfer-function of a minimum-size CMOS inverter indicates a rather large open-loop gain 'A' (tangent to the curve when $V_{IN} = V_{OUT}$). Both the transistors in the CMOS inverter are in SATURATION when $V_{IN} = V_{OUT}$. By using the drain-to-source path of a pMOS transistor in the feedback path and forcing its gate voltage to zero ($V_{TUNE} = 0$), the pMOS transistor appears to be a linear resistor about the operating point ($V_{IN} = V_{OUT}$) [figure 4-44]. The feedback pMOS resistor is in its TRIODE region at equilibrium because $V_{SG} \geq -V_{TP}$ ($2.5 \geq -(-1)$); a typical value for V_{TP} is $-1.0-V$) and $V_{SD} \leq V_{SG} + V_{TP}$ ($0 \leq 2.5 + (-1)$). Since the TRIODE pMOS I_D-V_{SD} relationship is given by [Eqn. 4-7], the conductance, g_{SD} , of the pMOS feedback resistance can also be obtained [Eqn. 4-8].

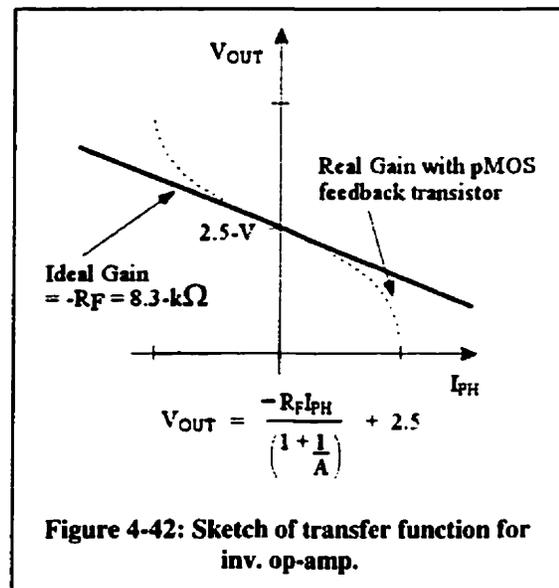


Figure 4-42: Sketch of transfer function for inv. op-amp.

Eqn. 4-7: $I_D = K_P[2(V_{SG} + V_{TP}) V_{SD} - V_{SD}^2]$

Eqn. 4-8: $g_{SD} = \delta I_D / \delta V_{SD} = K_P[2(V_{SG} + V_{TP}) - 2V_{SD}]$

(with $V_{SG} = 2.5\text{-V}$, $V_{SD} = 0\text{-V}$, $V_{TP} = -1.0\text{-V}$, and $K_P = 50\text{-}\mu\text{A/V}^2$)

therefore $g_{SD} = 1.5 \times 10^{-4} \text{ A/V}$ (or $6.67\text{-k}\Omega$)

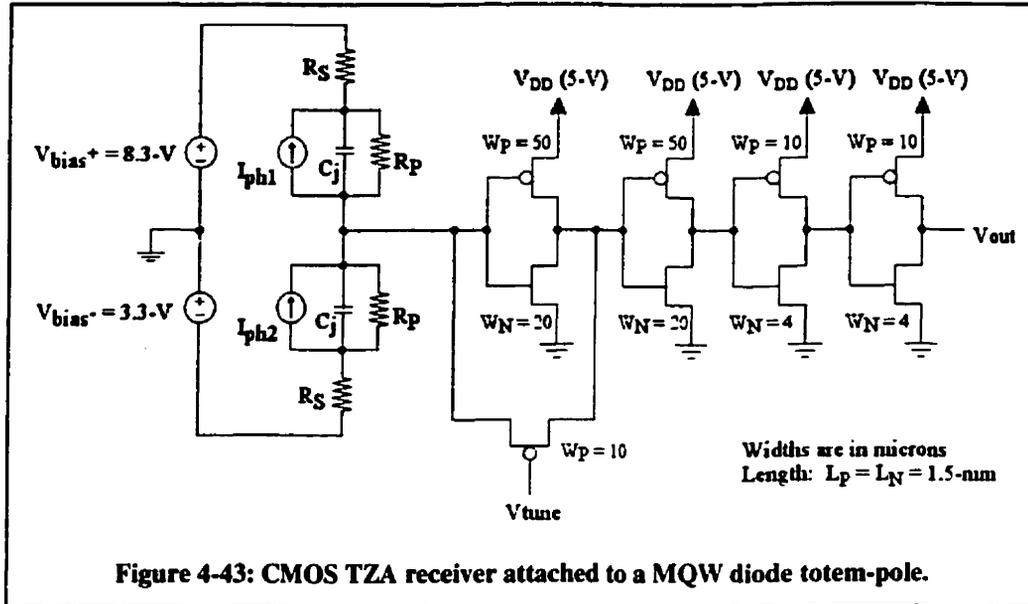


Figure 4-43: CMOS TZA receiver attached to a MQW diode totem-pole.

When $V_{IN} = V_{OUT}$, the center-tap point of the detector MQW diode totem-pole remains at 2.5-V . Using bias supply voltages of $V_{det+} = 8.3\text{-V}$ and $V_{det-} = -3.3\text{-V}$ the reverse bias voltage across each MQW diode can be maintained at $V_{MQW} = 5.8\text{-V}$ due to the constant voltage of 2.5-V at the input to the TZA. This voltage can remain virtually constant (at 2.5-V) due to the relatively high gain across the front-end TZA CMOS receiver. This allows the reverse-bias across the detecting MQW diodes to remain virtually constant.

Using the small-signal model for CMOS inverter in saturation and a similar model for the conductance of the feedback pMOS resistor, the dynamic

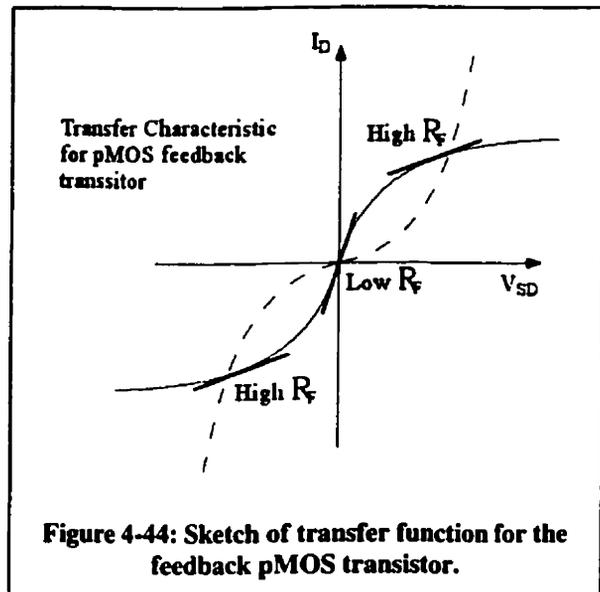


Figure 4-44: Sketch of transfer function for the feedback pMOS transistor.

operation of the transimpedance amplifier can be analyzed [figure 4-45] [Eqn. 4-9, 4-10, 4-11]. Using the simplified model, the following relationships for input and output resistance as well as the closed-loop transfer-function can be developed:

Eqn. 4-9: $R_{IN} = (R_f + r_o)/(1 - A)$

Eqn. 4-10: $R_{OUT} = r_o/(1 - A)$

Eqn. 4-11: $V_{OUT}/I_{IN} = - (R_f - r_o/A)/(1 - 1/A)$

Using the same procedure outlined in the previous section, the open-loop gain of a symmetric CMOS inverter-amplifier (when $W_P = 2.5W_N$) is obtained using the DC-

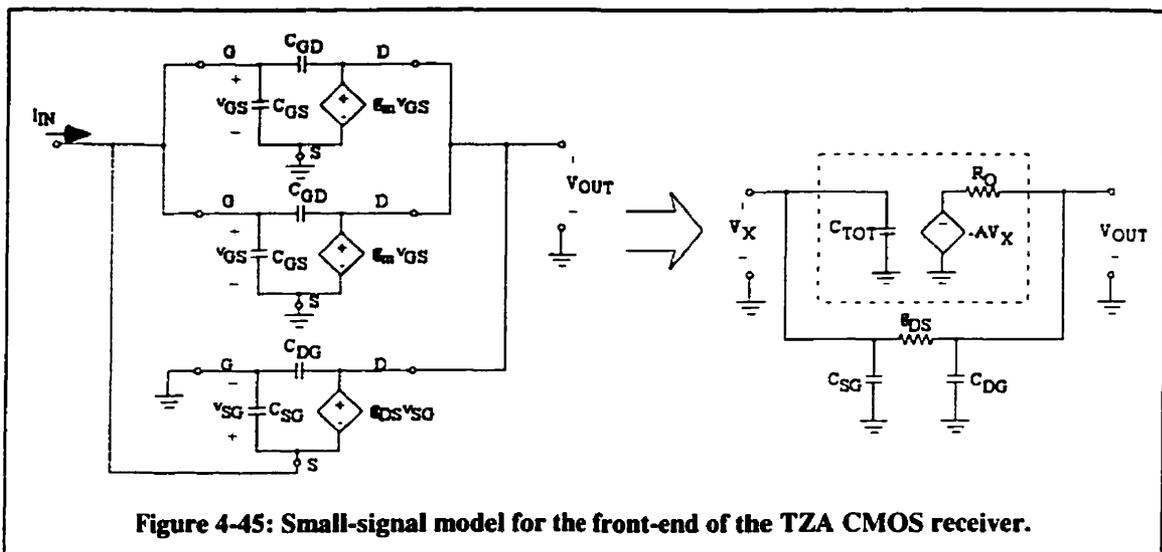


Figure 4-45: Small-signal model for the front-end of the TZA CMOS receiver.

sweep function in HSpice ($A = -25.77\text{-V/V}$). The values of the small-signal transconductance can also be obtained at the operating point from the output of the HSpice simulation (when $V_{IN} = V_{OUT} = 2.53\text{-V}$) $g_{mP} = 15.83\text{-}\mu\text{A/V}$ and $g_{mN} = 8.11\text{-}\mu\text{A/V}$. Since the controlling voltage, V_{IN} , is across the controlled voltage sources, the nMOS and pMOS transconductances become the total output resistance; $r_o = 1/(g_{mP} + g_{mN})$, hence $r_o = 41,776.85\text{-}\Omega$. The feedback pMOS transistor's small-signal model neglects the dependence of V_{SG} , and is linearized about the operating point. The feedback conductance obtained was $g_{SD} = 119.8\text{-}\mu\text{A/V}$ (or $R_f = 8.93\text{-k}\Omega$) (again these values are obtained from the output file of HSpice). The total input capacitance is obtained using the formula described earlier that involved the Miller-effect gain across the gate-to-drain

capacitances for each transistor. This capacitance is added to the input capacitance of the feedback pMOS transistor; $(C_{SG})_F$ [Eqn. 4-12]:

$$\begin{aligned}
 \text{Eqn. 4-12:} \quad C_{\text{TOT-AMP}} &= (C_{SG})_F + (C_G)_N + (C_G)_P + (1-A)((C_{DG})_N + (C_{GD})_P) \\
 &= 4.38\text{-fF} + 13.86\text{-fF} + 5.45\text{-fF} \\
 &\quad + (1 - (-25.77))(3.11\text{-fF} + 1.32\text{-fF}) \\
 &= 138.98\text{-fF}
 \end{aligned}$$

This capacitance, $C_{\text{TOT-AMP}}$, represents the total capacitance at the input to the transimpedance amplifier seen by the MQW photodiodes. To estimate the rise-time at the input node of the transimpedance amplifier, the capacitance of the MQW diodes must also be added. Assuming 500-fF for each MQW diode, the total capacitance would be [Eqn. 4-13]:

$$\text{Eqn. 4-13:} \quad C_{\text{TOT-MQW}} = 500\text{-fF} + 500\text{-fF} + C_{\text{TOT-AMP}} = 1138.98\text{-fF}$$

The total resistance [Eqn. 4-14] at the input node of the transimpedance amplifier is simply the input resistance, R_{IN} , given in the equation above [Eqn. 4-9]. The MQW diode resistances can be neglected since they are in parallel and greater than 10-M Ω .

$$\text{Eqn. 4-14:} \quad R_{\text{IN}} = (8.93\text{-k}\Omega + 41.78\text{-k}\Omega)/(1 - (-25.77)) = 1.89\text{-k}\Omega$$

Using the formula for the rise-time of a first-order circuit [Eqn. 4-15], the rise-time of the signal at the input node of the transimpedance amplifier can be estimated.

$$\text{Eqn. 4-15:} \quad t_{\text{rise}} = R_{\text{IN}}C_{\text{IN}}\ln(9) = (1.89\text{-k}\Omega)(1139\text{-fF})(\ln(9)) = 4.74\text{-ns}$$

When the entire receiver circuit was simulated using HSpice, the input node of the transimpedance amplifier had a rise-time of 4.66-ns. The rise-time of a minimum-sized inverter in the same technology was approximately 0.35-ns. This speed difference indicates the trade-off between absolute gain and bandwidth. Since a small photocurrent

was generated by the MQW photodiodes, the ideal input resistance of the amplifier would be a low input resistance. This is reasonably well satisfied by using the TZA with the feedback pMOS transistor.

By slightly increasing the voltage, V_{TUNE} , of the feedback pMOS resistor, the feedback resistance can be increased. This increases the sensitivity of the amplifier, but reduces the bandwidth. As the closed-loop gain, V_{OUT}/I_{PH} , increases, so does the input resistance and therefore so does the rise-time. Although the TZA can be made more sensitive, the bandwidth decreases until the pMOS feedback is broken and the open-loop gain is reached.

To obtain the previously stated total gain of 1,000,000 V/A, the TZA must be followed by several voltage gain stages. These stages are symmetric minimum-sized inverters, each with a gain of approximately -25.77 V/V each. Therefore, to achieve a gain of at least 1,000,000 V/A, the TZA must be followed by at least 2 inverter stages [Eqn. 4-16]:

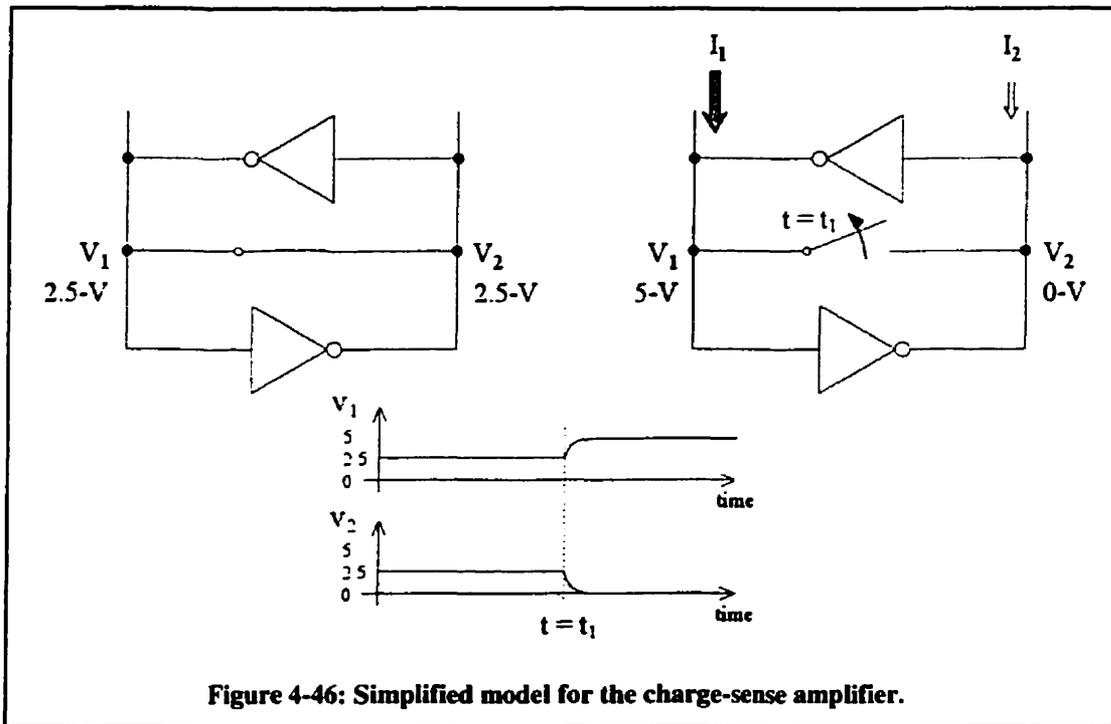
$$\begin{aligned} \text{Eqn. 4-16:} \quad \text{Gain}_{TOT} &= (-8930 \text{ V/A})(-25.77 \text{ V/V})(-25.77 \text{ V/V}) \\ &= -5,900,000 \text{ V/A} \end{aligned}$$

It is important that each stage in the amplifying circuit have the same threshold voltage. This is the voltage at which $V_{IN} = V_{OUT} = 2.53\text{-V}$. This condition is especially important at the input node of the TZA. However, this condition is easily satisfied by using the negative feedback pMOS transistor. The feedback forces the TZA to use the operating point ($V_{IN} = V_{OUT}$) as the voltage threshold.

One of the major problems experienced when using the TZA was the possibility of a floating-voltage at its input node. If no light was incident on the MQW photodiodes, the voltage at the input node would remain at +2.5-V. And, if all the subsequent CMOS gates were symmetric about +2.5-V, an unknown state could propagate through the chip. This problem was avoided using slightly asymmetric inverters after the TZA, so that an input voltage of +2.5-V would eventually cause one of the subsequent inverters to rail to a logic level. The floating-voltage problem was first observed in the testing of the Workshop chip (see Chapter 5 - Experimental Results).

4.6.6.b) The clocked-charge sense amplifier

The clocked-charge sense amplifier (C-CSA) was the optical receiver design used in the Phase-III-A (P3A) chip. The designers for this circuit were Dr. Alain Shang (now at Honeywell), and Mr. Pritam Sinha (now at ATI Technologies Inc.). The circuit was used because it could be designed in a small area and have an extremely low static power consumption. Also, it could be made very sensitive to low currents and thus relatively independent to the input capacitive load. In theory, any size MQW diode could be attached to the C-CSA receiver without affecting its performance. The C-CSA was also by nature a symmetric circuit and required both the signal and its complement to sense



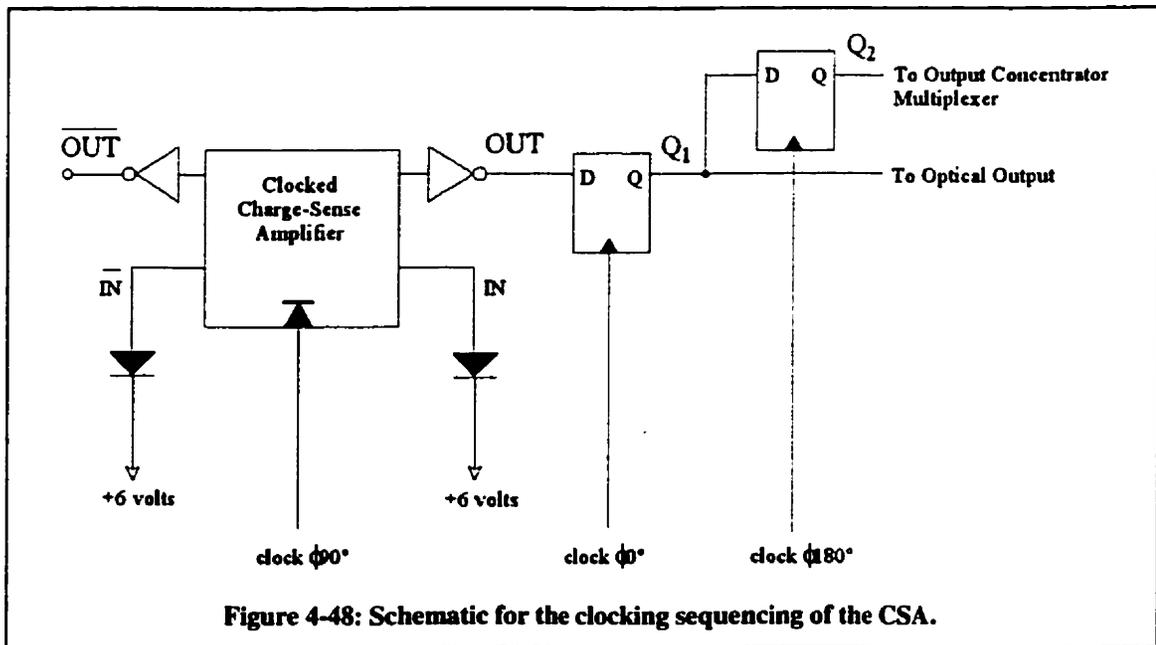
the data; similar to the dual-rail encoding used in all the systems. Although most of the analysis of this circuit can be found elsewhere [ref 22], a brief explanation of the C-CSA is given here, as well as some techniques used in the layout of the C-CSA.

The C-CSA has been typically used in large memory arrays [ref 23] to decrease the power required to sense and amplify the state of the memory cells. The operation of this circuit is based on a pair of identical inverters connected in positive feedback. A

function of inverter Mn1/Mp1 would have shifted more than the inverter Mn2/Mp2. The output node, O2, would be driven low and the output node, O1, would be driven high.

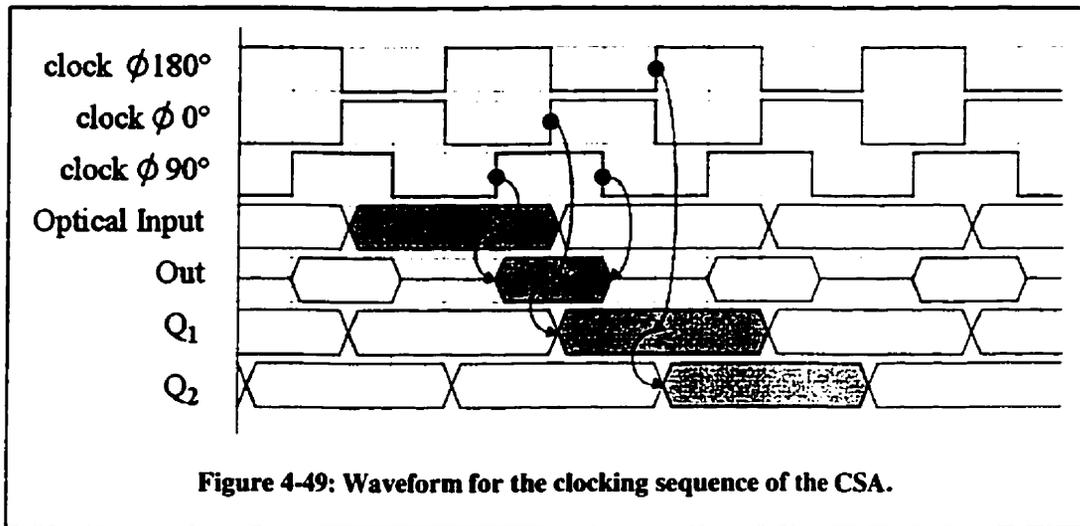
Since the C-CSA required a signal to reset its state and place the inverter pair back into a meta-stable state, the C-CSA was considered as a level sensitive latch or delay element. The clock supplied to the transistors Mp3 and Mp4 caused an interval of meta-stability when high and an interval of data when low. Due to the meta-stable interval, the C-CSA had to be followed by an edge-triggered D-flip flop (DFF) that could sample the C-CSA's output in the middle of the data interval. The DFF would prevent the meta-stable state from propagating further into the chip that could cause undesired effects in the CMOS circuitry. The C-CSA clock and the DFF clock were therefore related by a phase-shift of a quarter of a period but with the same frequency [figure 4-48] [figure 4-49]. This required an additional clock to be routed into the chip for each C-CSA in the array.

The most critical design requirement of the C-CSA was that the circuit was built using matched devices. Even a slight mismatch could induce a bias towards one output



state. The greater the mismatch, the larger the differential optical power required to operate correctly. These errors are very much fabrication dependant and can only be minimized for during the design using specific design rules. The drawn size of the

inverters can be relatively well matched, but the doping variations during fabrication can cause mismatches in threshold voltage and resistance. Therefore, the two inverters are laid-out as close as possible to one another to decrease the variation. The loading on the inverter pair also had to be equal. Two identically sized inverters, Mn6/Mp6 and Mn7/Mp7, were used as the output loads for the inverter pair. However, the digital logic



required only one output from the C-CSA and thus only the inverter Mn6/Mp6 had its output lead to the smart pixel circuit. The output of the other inverter Mn7/Mp7 was left floating.

A final design requirement was to isolate the C-CSA from any substrate or carrier-injected noise. Several isolation rings were laid-out around each C-CSA. Since the circuit was surrounded by p⁻ substrate, a ring of p⁺-diffusion, attached to the power supply, was used to enclose this region. This mechanism allowed a 'sweep-out' any minority carriers in the substrate immediately around the C-CSA. A second ring of n⁺-diffusion in an n-well was connected to the ground supply and was used to isolate the C-CSA circuits from one another. The n-well region is typically diffused approximately 15-μm into the substrate and therefore can block most lateral movement of minority carriers. The isolation rings were also used to block any substrate photocarriers generated by stray optical beams on the surface of the silicon. Although one of the layout strategies in the P3A design was to keep electronic circuitry away from the MQW diode clusters, the C-CSA were placed below the MQW diodes. It was believed that the C-CSA circuitry

would be disturbed far more by the smart pixel digital switching circuits than stray photocarriers generated by a few microWatts of optical power. Also, the C-CSA circuitry was placed directly below the rather large MQW diodes that would have blocked virtually all the light from reaching the C-CSA circuit.

4.7) Conclusion

This chapter outlined the design of four VLSI-OE chips. Although the chips presented in this chapter were based on the same basic architecture, there were many layout strategies and testing strategies that were developed in this work.

The most important aspect of the design of these chips was the relationship between the layout of the electronic circuitry and the design of the optical relays. This relationship manifested itself in the arrangement of the optoelectronic devices on the surface of the VLSI-OE chips. A compromise between the ideal circuit layout and the ideal optical layout was implemented using a matrix of clustered arrays of optoelectronics. Due to the nature of the clustered optoelectronics, a novel structure called the super-cluster was developed that allowed the proper logic functions to be implemented between a cluster of detector devices and a cluster of modulator devices. The key to this layout technique was that the design of the super-cluster took advantage of symmetrical layout and was created in a manner that allowed it to be repeatable and abutted side-by-side. This layout technique allowed more flexibility in the optical design.

Another characteristic concerning the design of the VLSI-OE chips was the ratio between the number of optical i/o to the number of electrical i/o. The type of architecture used for the optical backplane was essentially a firehose-architecture. This meant that the amount of electrical data passing into and out of a chip was only a fraction of the optical data passing through a chip. Methods such as address recognition as well as efficient methods to route optical input data to the electrical output, such as the output concentrator tree, were required. Other techniques, such as memory buffering, could be used but these were not explored.

The detailed analysis of the performance of certain optoelectronic receivers, such as the transimpedance amplifier and the charge-sense amplifier, helped to define the physical layout requirements by defining the performance and power requirements as a function of the receiver's design parameters. The analysis of the driver circuitry also helped determine layout requirements by providing estimates of the power consumption and sizes for the transistors.

In the next chapter, the results of tests performed on each of these VLSI-OE chips will be provided. The tests will describe speed as well as functional testing and will also involve some test results which indicate specific design problems and which lead to specific changes in the design of the next iteration of VLSI-OE chip.

4.8) References:

- [1] A.V. Krishnamoorthy, D.A.B. Miller, "Scaling optoelectronic-VLSI circuits into the 21st century: a technology roadmap", *IEEE Journal on Selected Topics in Quantum Electronics*, Vol. 2, No. 1, April 1996, pp. 55-76
- [2] A.V. Krishnamoorthy, K.W. Goossen, "Progress in optoelectronic-VLSI smart pixel technology based on GaAs/AlGaAs MQW modulators", *International Journal of Optoelectronics*, Vol. 11, No. 30, 1997, pp. 181-198.
- [3] T.K. Woodward, A.V. Krishnamoorthy, A.L. Lentine, L.M.F. Chirovsky, "Optical receivers for optoelectronic VLSI", *IEEE Journal on Selected Topics in Quantum Electronics*, Vol. 2, No. 1, April 1996, pp. 106-116
- [4] D.R. Rolston, B. Robertson, H.S. Hinton, D.V. Plant, "Analysis of a microchannel interconnect based on the clustering of smart pixel device windows", *Applied Optics*, Vol. 35, No. 8, March 1996, pp. 1220-1233.
- [5] T. Szymanski, H.S. Hinton, Optoelectronic smart pixel array for a reconfigurable intelligent optical backplane, United States Patent # 6,016,211, Issued Jan 18, 2000.
- [6] T.H. Szymanski, V. Tyan, "Error and flow control on terabit intelligent optical backplanes" *IEEE Journal on Selected Topics in Quantum Electronics*, Vol. 5, No. 2, March-April 1999, pp. 339-352
- [7] D. Kabal, Packaging of surface active optoelectronic device arrays, Master's Thesis, McGill University, Montreal, Canada, 1997
- [8] P. Pavan, G. Spiazzi, E. Zanoni, M. Muschitiello, M. Cecchetti, "Latch-up DC triggering and holding characteristics of n-well, twin-tub and epitaxial CMOS technologies Circuits", *IEE Proceedings of Devices and Systems*, Vol. 138, No. 5, Oct. 1991, pp. 604-612.
- [9] B. Robertson, Y. Liu, G.C. Boisset, M.R. Tagizadeh, D.V. Plant, "In situ interferometric alignment systems for the assembly of microchannel relay systems", *Applied Optics*, Vol. 36, No. 35, Dec. 1997, pp. 9253-9260.
- [10] M. Ayliffe, D.V. Plant, "On the design of misalignment-tolerant free-space optical interconnects", *Optics in Computing 2000*, SPIE Vol. 4089, Quebec City, Canada, 2000, pp. 905-916.
- [11] M.H. Ayliffe, D.R. Rolston, E.L. Chuah, E. Bernier, F.S.J. Michael, D. Kabal, A.G. Kirk, D.V. Plant, "Packaging of an optoelectronic-VLSI chip supporting a 32x32 array of surface-active devices", *Optics in Computing 2000*, SPIE Vol. 4089, Quebec City, Canada, 2000, pp. 508-518.
- [12] D.A.B. Miller, D.S. Chemla, S. Schmitt-Rink, "Electroabsorption of highly confined systems: Theory of quantum-confined Franz-Keldysh effect in semiconductor wires and dots" *Applied Physics Letters*, Vol. 52, 1988, pp. 2154-2156.
- [13] A.L. Lentine, D.A.B. Miller, "Evolution of the SEED technology: bistable logic gates to optoelectronic smart pixels", *IEEE Journal of Quantum Electronics*, Vol. 29, No. 2, Feb. 1993, pp. 655-669
- [14] D.A.B. Miller, "Novel analog self-electro-optic-effect devices", *IEEE Journal of Quantum Electronics*, Vol. 29, No. 2, Feb. 1993, pp. 678-698.
- [15] M. Tamamura, S. Shiotsu, M. Hojo, K. Nomura, S. Emori, H. Ichikawa, T. Akai, "A 9.5-Gb/s Si-bipolar ECL array", *IEEE Journal of Solid-State Circuits*, Vol. 27, No. 11, Nov. 1992, pp. 1575-1578

- [16] M.B. Venditti, D.N. Kabal, M.H. Ayliffe, D.V. Plant, F.A.P. Tooley, E. Richard, J. Currie, A.J. Spring Thorpes, "Temperature dependence of QCSE device characteristics and performance", IEEE/LEOS Summer Topical Meetings, 1998, pp. IV/17 -IV/20
- [17] D.B. Buchholz, A.L. Lentine, R.A. Novotny, "Thermal shift in the exciton absorption maxima as a function of the chip package design", Proceedings of the SPIE - The International Society for Optical Engineering, San Jose, CA, Feb. 1996, Vol. 2691, pp. 162-170.
- [18] R.A. Novotny, "Analysis of Smart Pixel Digital Logic and Optical Interconnects", Ph.D. Thesis, Heriot-Watt University, Edinburgh, Scotland, 1996.
- [19] N.H.E. Weste, K. Eshraghian, Principles of CMOS VLSI Design 2nd Ed., Addison-Wessley, New York, 1992
- [20] T.K. Woodward, A.V. Krishnamoorthy, A.L. Lentine, K.W. Goossen, J.A. Walker, J.E. Cunningham, W.Y. Jan, L.A. D'Asaro, L.M.F. Chirovsky, S.P. Hui, B. Tseng, D. Kossives, D. Dahringer, R.E. Leibenguth, "1-Gb/s two-beam transimpedance smart-pixel optical receivers made from hybrid GaAs MQW modulators bonded to 0.8-micron silicon CMOS", IEEE Photonics Technology Letters, Vol. 8, No. 3, March 1996, pp. 422-424.
- [21] A.S. Sedra, K.C. Smith, Microelectronic Circuits 4th Ed., Oxford University Press, New York, 1997.
- [22] T.K. Woodward, A.V. Krishnamoorthy, K.W. Goossen, J.A. Walker, J.E. Cunningham, W.Y. Jan, L.M.F. Chirovsky, S.P. Hui, B. Tseng, D. Kossives, D. Dahringer, D. Bacon, R.E. Leibenguth, "Clocked-sense amplifier-based smart pixel optical receivers", IEEE Photonics technology Letters, Vol. 8, No. 8, Aug. 1996, pp. 1067-1069.
- [23] E. Seevinck, P.J. van Beers, H. Ontop, "Current-mode techniques for high-speed VLSI circuits with application to current sense amplifier for CMOS SRAMs", IEEE Journal of Solid State Circuits, Vol. 26, No. 4, Apr. 1991, pp. 525-536.

Chapter 5: Experimental Results

5.1) Introduction

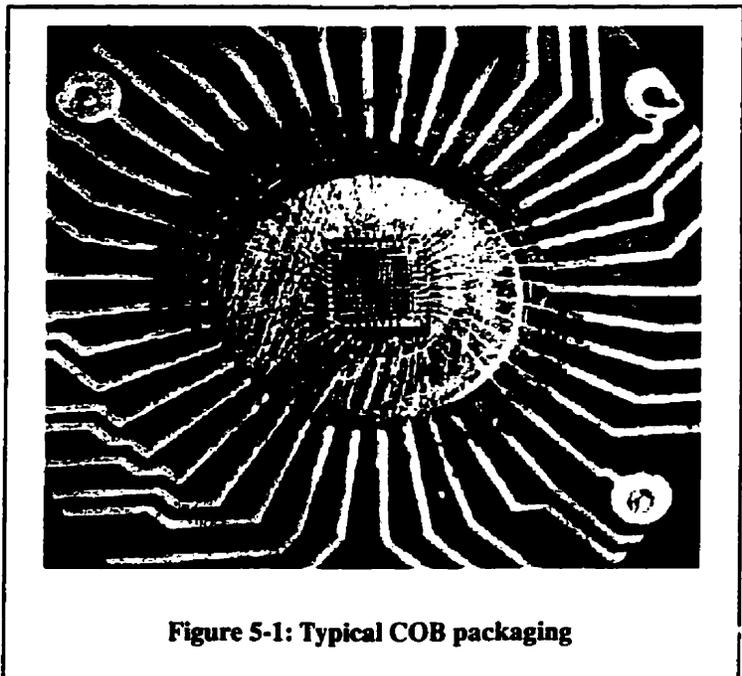
This chapter will first briefly describe the type of packaging and the support electronics required to operate the VLSI-OE chips. Data related to the performance and initial testing of several VLSI-OE chip will then be provided. The Phase-III-B chip will be described in slightly more detail than the other chips, since its functionality and operation were the most optimal. Results such as maximum bit rate, rise-times and fall-times will be described.

5.2) Packaging and external control

Each of the four chips required a specialized package that provided several features essential to chip operation. The package had to provide mechanical support for the chip as well as mechanical isolation from the support electronics via a flexible connector. The package required a method of thermal management, and it also required a method of supplying all the signals and power supplies to the chip. Furthermore, the package had to be of the proper size so that it was possible to integrate it into the optical system. The VLSI-OE chip package had to maintain a precise alignment between the optical relay and the surface of the chip. In the Phase-II system, the chips had to be placed 880- μm away from their respective lens arrays. The Phase-III system required an 8-mm separation between the chips and the lens arrays, but had a very strict tolerance on tilt. High-speed electrical connections between the VLSI-OE chip and external circuitry were also required to demonstrate the speed at which electrical-to-optical and optical-to-electrical communications could occur. Unfortunately, traditional off-the-shelf packages could not satisfy these requirements very easily. Off-the-shelf packages were either too physically bulky or had relatively poor on/off-chip performance in terms of their transmission line effects. However, the most difficult obstacle with an off-the-shelf package was that it caused uncertainty in the chip to lens alignment. A chip package soldered onto a PCB could be placed to within tens of microns, but not to within ± 5

microns. This uncertainty in position tolerance would have required complicated external mechanical positioning techniques. Therefore, the technique of chip-on-board (COB) mounting [ref 1] was used so that fewer sources of misalignment were present and so that the electrical performance of the package was optimized. Using COB mounting, a VLSI-OE chip could be placed next to the trace lines on a PCB and wire-bonded to the fingers of the PCB. This technique eliminates the need for pins and solder joints of most chip packages allowing for only one-level of trace-line discontinuity [figure 5-1]. The COB packaging was accomplished using relatively standard PCB design tools and fabrication techniques. Another useful feature of the COB package was that the fingers on the PCB could be tailored to the bond-pad layout on the VLSI-OE chip.

A crucial aspect of each of the designs was the method of mechanical isolation between the VLSI-OE chip and the rest of the support electronics. The chip and optical system were fixed together, and the flexible ribbon connector or the flexible PCB allowed the external control electronics to be



manipulated without affecting the chip to lens alignment. In the Phase-II design, the Beta-Chip and Workshop-Chip were placed on a rigid COB PCB with a high-speed strip-line flexible ribbon connector. The primary designers of this connector were Dr. Guillaume Boisset (now at Corning in Corning, N.Y.) and Mr. David Kabal (now at Nortel Networks in Kanata, Ontario). The COB PCB was called the Daughter-board [figure 5-2] [ref 2]. The Phase-III design used a slightly different technique involving flexible PCBs and more aggressive techniques for thermal management. Almost 200 signals and bias lines were

required on the Phase-III design [figure 5-3] [ref 3]. The lead designers for the Phase-III Flex-PCB designs were Mr. David Kabal, Mr. Michael Ayliffe and Mr. Alan Chuah.

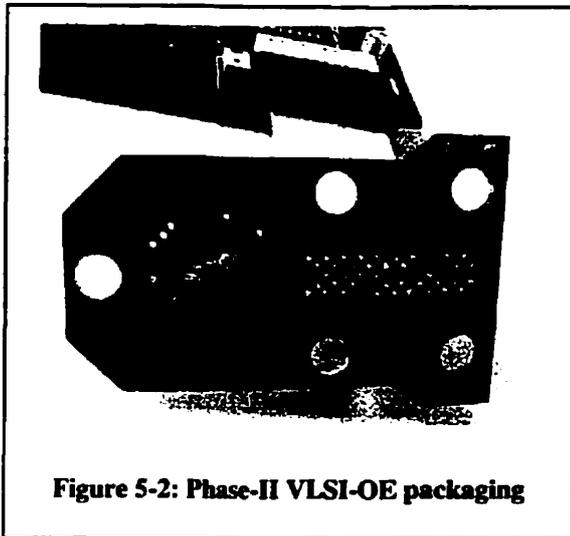


Figure 5-2: Phase-II VLSI-OE packaging

External support electronics was designed for each of the four chips. The external electronics designed and implemented by the McGill Photonic Systems group was not part of the Hyperplane design and was used strictly to test the functional behavior and determine the operating speed of the chips. The Beta-Chip and Workshop-Chip used an external Mother-Board to implement on-board

counters and methods to control the smart pixels on the chip. The Phase-III-A and Phase-III-B chips were also tested using external mother-boards that had on-board counters and

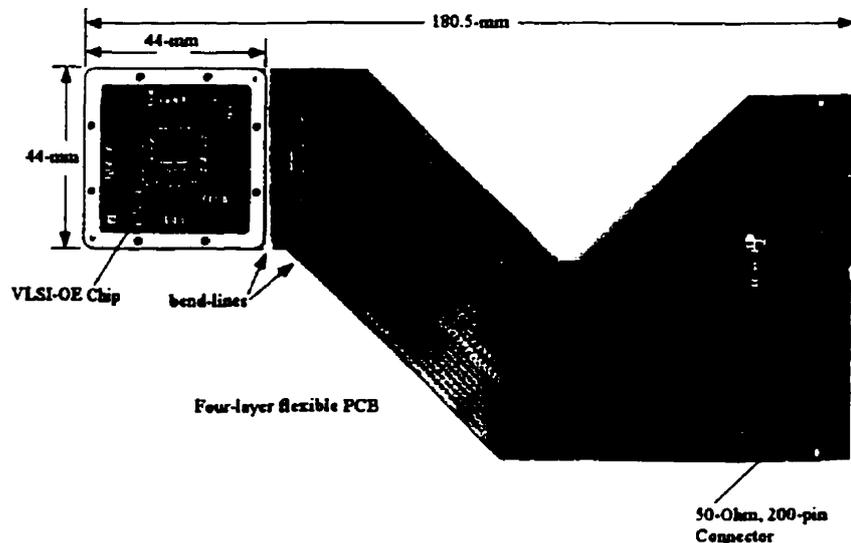


Figure 5-3: Phase-III flex-PCB packaging

LFSR circuitry that could generate pseudo-random bit patterns. The Phase-III motherboards could also interfaced with custom software on a PC so that the boards could be configured without requiring direct access to the Mother-Board [figure 5-4]. The Phase-III motherboards were designed by several people within the group, but the lead

designers of the P3A Mother-Board were Ms. Madeleine Mony, Ms. Emmanuelle Laprise and Mr. Michael Venditti and the lead designers of the P3B motherboard were Mr. Feras Michael and Mr. Alan Chuah.

5.3) Simulation and experimental results

5.3.1) The Beta-Chip

The design of the Beta-Chip was simulated using level-3 HSpice transistors models. The simulations did not include parasitic effects of internal capacitance or resistance due to the large simulation times involved. However, the simulations did

verify the functional operation of the chip before it was fabricated. Although most of the circuit simulations were performed using 100-Mbps data rates, these rates could not accurately reflect the true data rates of a fabricated chip.

Once the Beta-Chip had been fabricated, the silicon digital circuitry was tested before the optoelectronics were attached to the surface of the silicon chip. The electrical Beta-Chip was packaged in a test-rig and placed under a microscope. Flexible cat-whisker micromanipulator probes, with 10- μm probe tips, were used under the microscope to simulate data into and out-of the optoelectronic

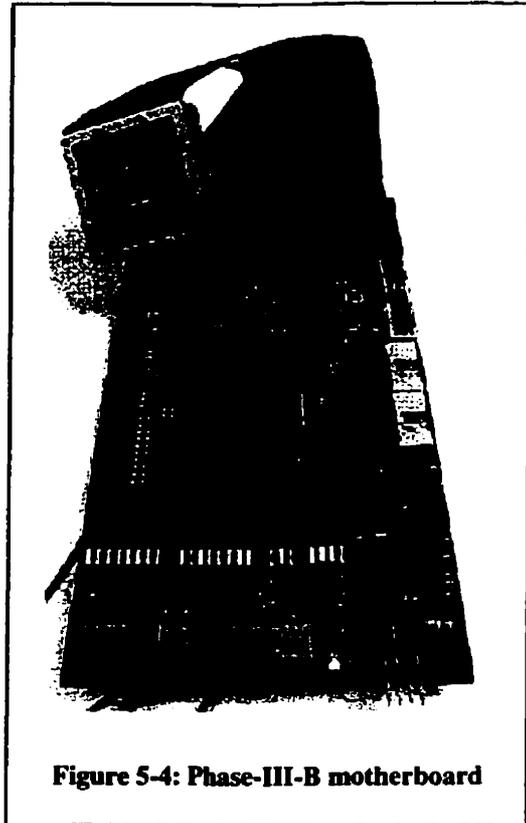


Figure 5-4: Phase-III-B motherboard

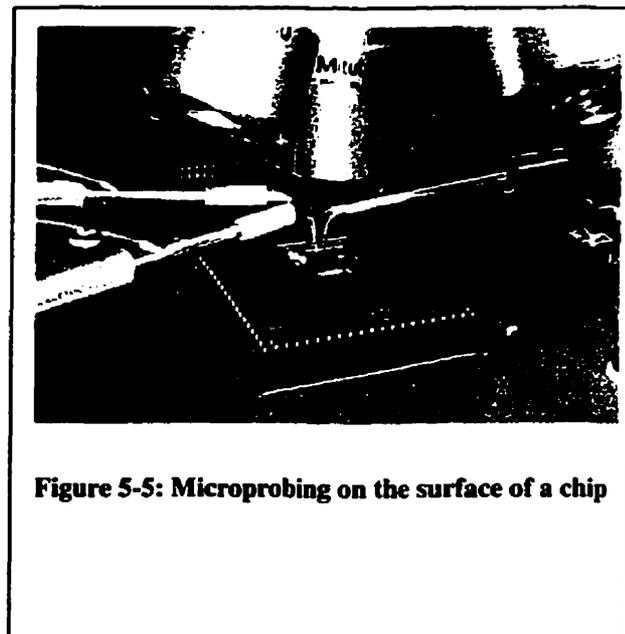


Figure 5-5: Microprobing on the surface of a chip

contact points. Two probes contacted the receiver totem-pole optoelectronics contact points to simulate the input photocurrents, and the third was used to read-out the state of the output transmitter driver [figure 5-5]. Low speed tests were performed to verify the functionality of the Bet-Chip.

To test the optical Beta-Chip, an eight-beam optical relay was used. The optical relay was based on a previous system demonstrator used to interconnect 2-PCBs [figure 5-6] [ref 4]. The optical relay used a baseplate approach and directed constant CW optical beams at one chip, which in turn modulated the beams with data and directed them to the receiver chip. Data rates of 1-Mbps, 10-Mbps, and 20-Mbps were obtained using this optical setup along with 16-bit patterns driven at a 1-MHz clock rate [ref 5] [figure 5-7]. The distorted bit patterns above 20-Mbps were due transmission line reflections between the output bond-pads and oscilloscope. Although these line reflections made it difficult to measure the performance of the Beta-Chip above 20-MHz, the rise-time and fall-time of the first part of the transition were significantly fast enough for 100-MHz operation. The 20-Mbps pattern was approximately 3.4-nsec.

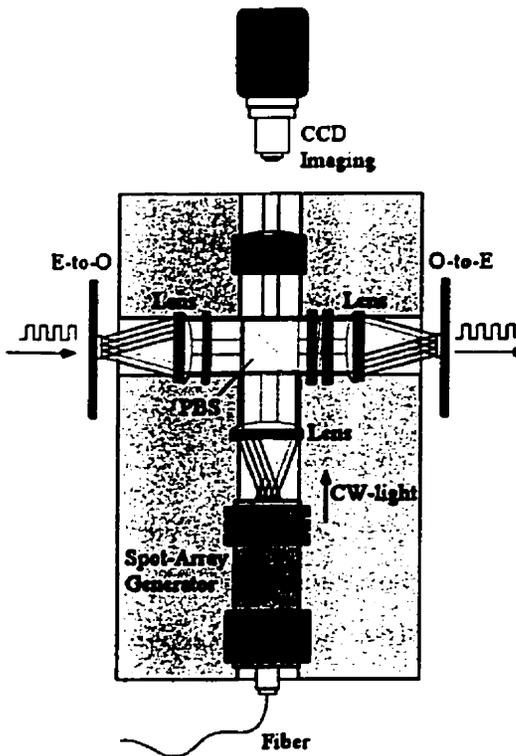


Figure 5-6: Schematic of optical test-rig for Beta-Chip

Unfortunately, no data was obtained for the Beta-Chip within the completed Phase-II system due to problems with the optoelectronics and the assembly of the final optical system.

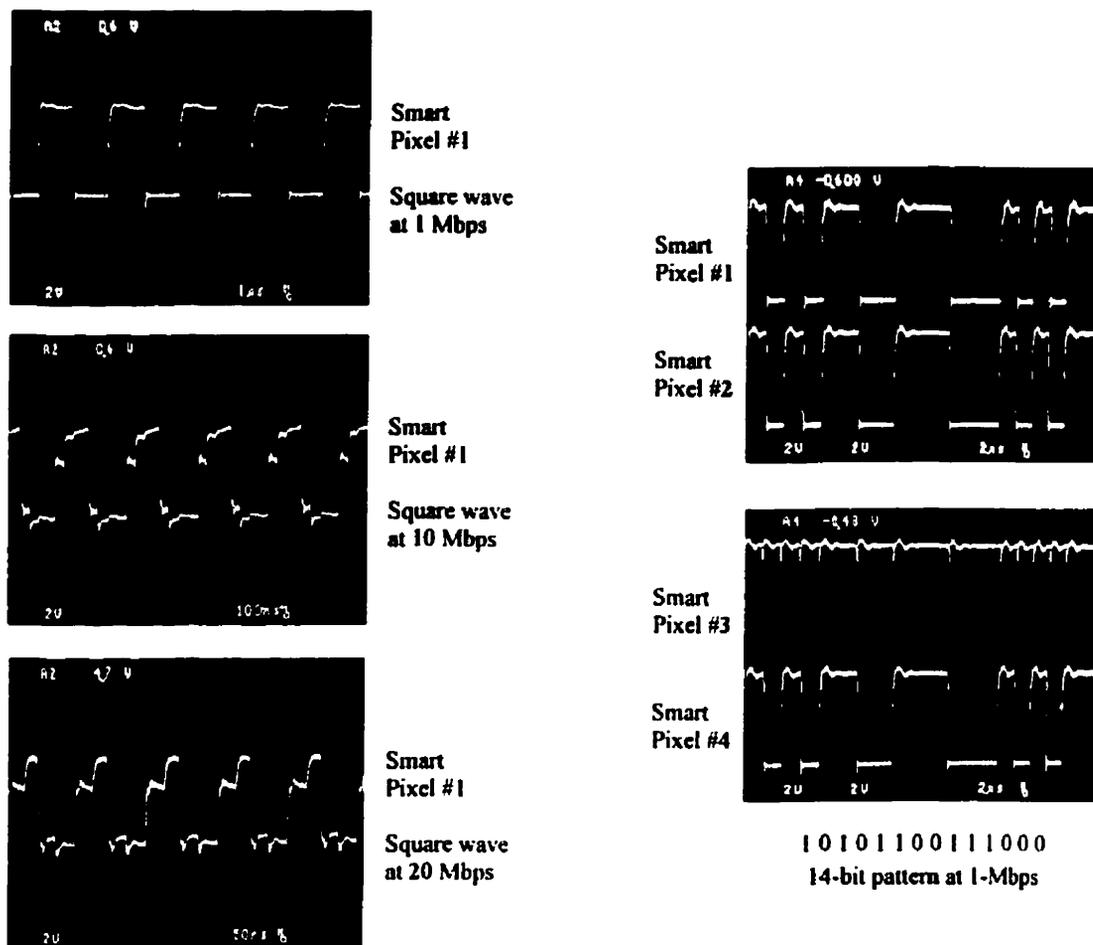


Figure 5-7: Data for Beta-Chip VLSI-OE

One of the main difficulties was the percent reflectivity of the Beta-Chip optoelectronic devices. The MQW diodes had roughly a 7-% and 12-% reflectivity in the low and high reflection states, respectively. The initial design required a worst-case of 15-% and 30-% for the low and high values because the optical system could provide only 100- μ W on the modulating MQW diodes. The poor reflectivity of the modulator diodes allowed only 3.5- μ W and 6- μ W to reach the detector MQW diodes at the next stage in the optical relay. The low differential power was not sufficient to switch the transimpedance amplifiers on the Beta-Chip within the Phase-II optical system. It was

only with powers above approximately 80- μ W that the Beta-Chip optical receivers could detect a signal. Further investigation of the performance of the Beta-Chip within the 2-PCB optical test set-up revealed that the detector MQW diodes were very sensitive to misalignment. If lateral misalignments of more than $\pm 3\text{-}\mu\text{m}$ from the center of the detectors were incurred, the signal would vanish.

5.3.2) *The Workshop-Chip*

The performance of the Workshop-Chip was extremely poor. Not only was the digital design poorly implemented, but the percentage reflectivities for the high and low states of the MQW diodes had an average of 6-% and 12-% for either state. These reflectivities were not useful in the Phase-II system.

Although the Workshop-Chip could not be used to obtain data on operational smart pixels, several interesting features of VLSI-OE chip operation were obtained. Prior to normal operation, the Workshop-Chip had to be configured using an internal 16-bit

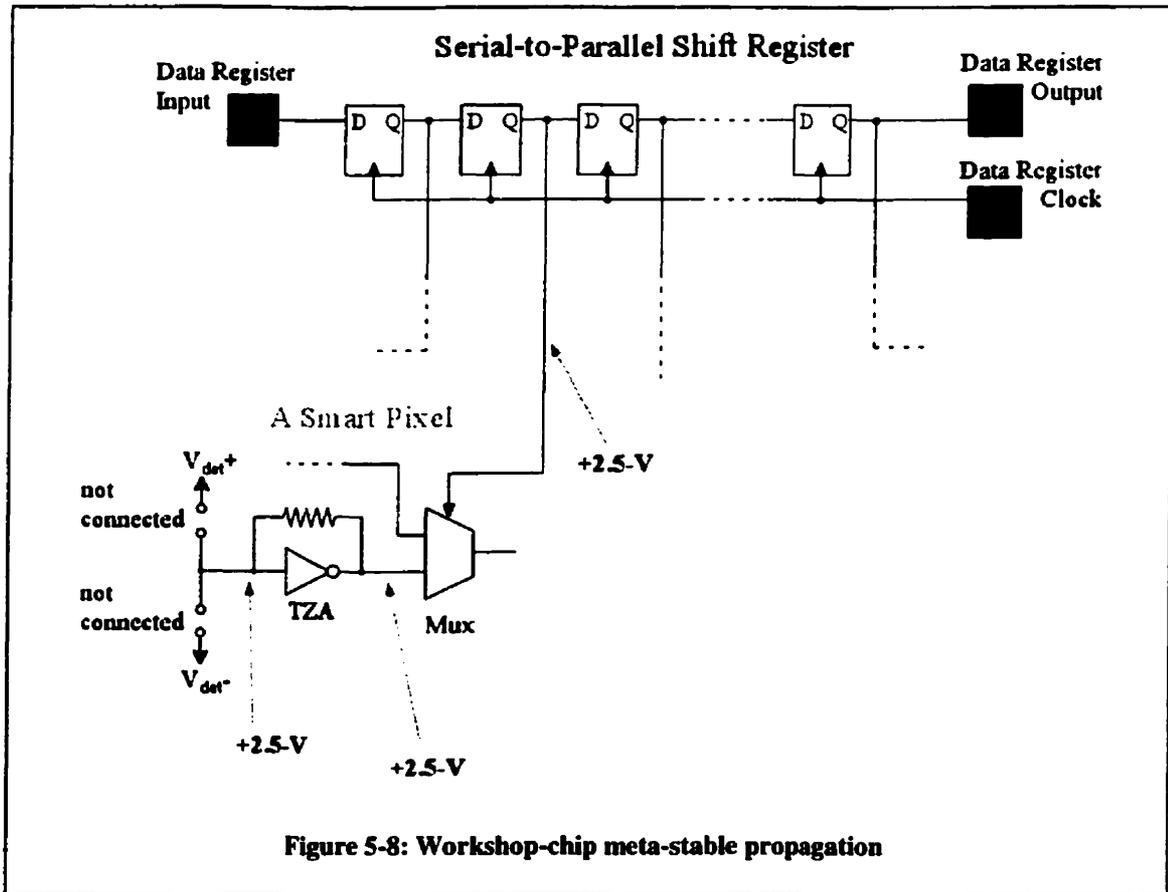


Figure 5-8: Workshop-chip meta-stable propagation

serial-to-parallel shift register. A bit pattern and a clock were supplied to the internal register which contained control bits for the smart pixel array. These control bits forced the smart pixels into known input-output states, and were primarily responsible for the correct routing of the data through the R-Muxes (see Chapter 4 – VLSI Optoelectronics). Unfortunately, it was observed that it was impossible to correctly set the state of the serial-to-parallel shift register. The control bits always seemed to be random and therefore the R-Muxes would never could be set properly. It was determined that the cause of this random behavior was related to the transimpedance (TZA) optical receivers used within each smart pixel and the type of pass-gate logic used to implement the multiplexers in the smart pixels. When the inputs to the TZAs were floating (i.e.: when no light was incident on the MQW diodes), the output of each TZA would produce a meta-stable voltage of approximately 2.5-V (between the 0-V and 5-V rails of the logic) [figure 5-8]. Since the TZAs were directly connected to pass-gate 2-to-1 multiplexers and the control signals of the multiplexers were connected to the serial-to-parallel register bits, the meta-stable states from the TZAs were affecting the serial-to-parallel register bits.

This conclusion was obtained by forward-biasing the detector MQW diodes. This was done by grounding both V_{det+} and V_{det-} . Since the node at the input of this type of TZA would naturally produce +2.5-V, one of the diodes in the totem-pole would turn-on and adjust the input node voltage to approximately 0.5-V [figure 5-9]. This was sufficient to force the output of the TZAs to a stable voltage. Once this was done, the serial-to-parallel shift register could be programmed properly. Unfortunately, the chip could no longer function as designed in this state.

The sensitivilty of pass-

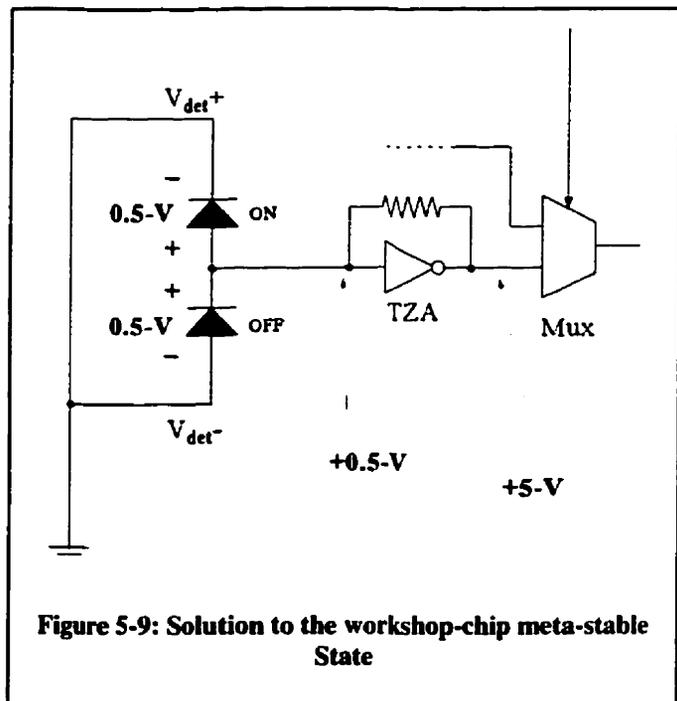


Figure 5-9: Solution to the workshop-chip meta-stable State

gate/CMOS designs to meta-stable states was exceptionally high in the Workshop-Chip design. The floating TZA input nodes were similar to having 16 floating input bond pads around the perimeter of the chip. The HCxx and HCTxx family of High-speed CMOS IC's available on the market today are very sensitive to floating inputs and it is typically suggested that any unused inputs be forced to either a low or high state and not left floating [ref 6]. Unfortunately, this was not known at the time of the design and the meta-stable states were able to propagate throughout the circuits and cause the serial-to-parallel register to behave poorly. Later designs used mechanisms such as the clocked charge-sense amplifier or asymmetric TZA amplifiers attached to more robust digital circuits to avoid the problems associated meta-stable voltage conditions.

5.3.3) The P3A Chip

For the initial testing of the Phase-III-A (P3A) chip, a VLSI-OE test-rig was constructed. A pin-grid-array (PGA) chip carrier with 256-pins and a top-open cavity was used to hold the P3A chip. It provided 2 tiers of bond fingers, the upper tier was for power and biasing, and the lower tier was for signal routing. The package was organized so that it could be used with a variety of chips with less than 100 signal bond pads. A custom designed PCB was also fabricated so that the 256-pin PGA could interface with a digital i/o card in a PC.

The PCB was laid-out and tested by Mr. Danny Birdie (now at Nortel Networks) and the test board was called the DB-Board. The DB-Board could be mounted under a microscope workstation

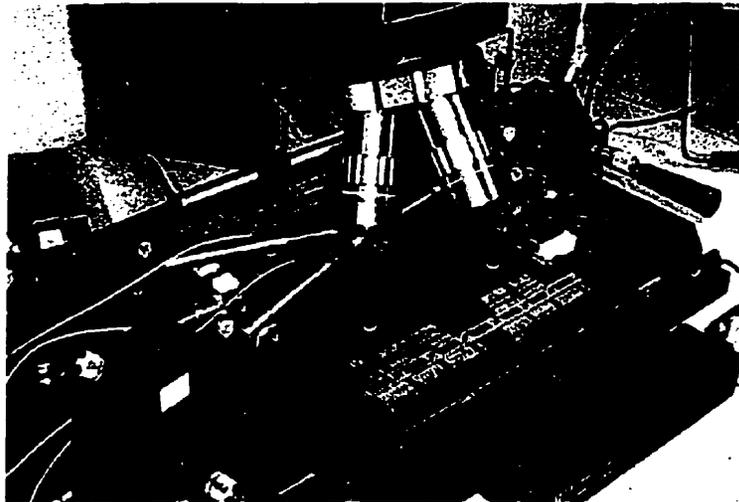


Figure 5-10: Microscope probing facility

and several micro-positioning probes could be used to access the top of the chip [figure 5-10]. The chip could be electrically controlled using a National Instruments 96-pin DIO boardTM that supplied a total of 96 data lines that could be configured in 12 banks of i/o bytes. Since the DB-Board was a generic testing platform, a set of VERO speed-wire pinsTM were placed between the 96-pin DIO connectors and the 256-pin PGA to allow the user to assemble a manual cross-connect using wire-wrap wires. This allowed complete flexibility as well as a generic platform to test other VLSI-OE chips. The PC software that accessed the 96-pin DIO board was custom designed by Ms. Emmanuelle Laprise, and allowed the user to interface with the VLSI-OE chip through a Windows driven display. The operating speed of the test bench was limited to less than 1-MHz by the wire-wrap connections.

The initial testing of the P3A chip indicated that most of the structures operated properly. Both the 28-bit serial-to-parallel registers functioned properly as well as the transmit address selection and transmit enable for the inject state of each channel. The transparent state of the smart pixels also worked as expected when two correctly phased clock inputs were used (see Chapter 4 – VLSI Optoelectronics). However, preliminary testing of the P3A chip in the DB-Board revealed three serious flaws in the overall design.

The first design flaw involved the extraction state of the channels. Essentially, the secondary D-FF within the smart pixel was enabled using an ACTIVE HIGH signal and the output concentrator multiplexer was enabled using an ACTIVE LOW signal. However, because of the logic structure used, either two ACTIVE HIGH signals or two ACTIVE LOW signals were sent. Therefore, there was no way to activate both mechanisms at the same time [figure 5-11]. The way this problem was overcome was to have chip microsurgery performed on the P3A chips. A company called Fibics Inc. in Ottawa, Ontario was able to sever metal trace lines and then join them to other metal trace lines for each channel in the chip. This effectively merged the control of the

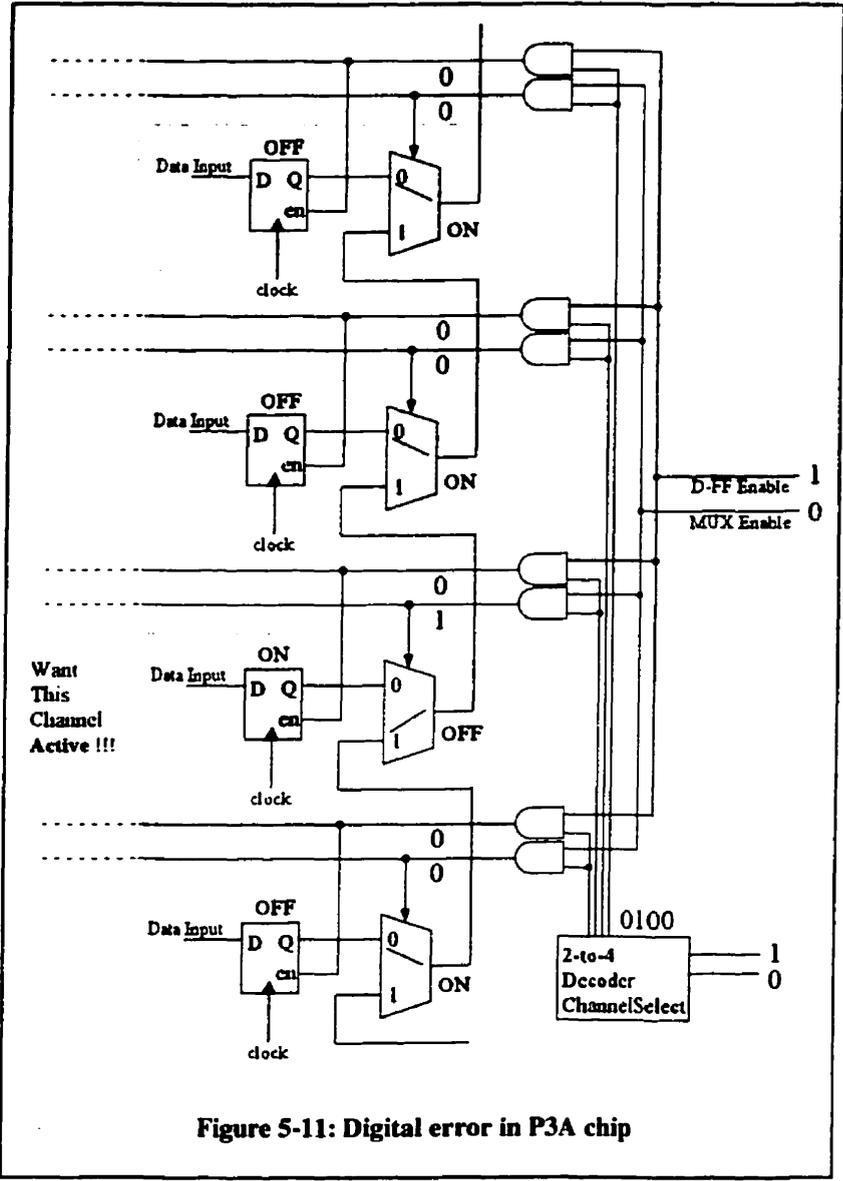
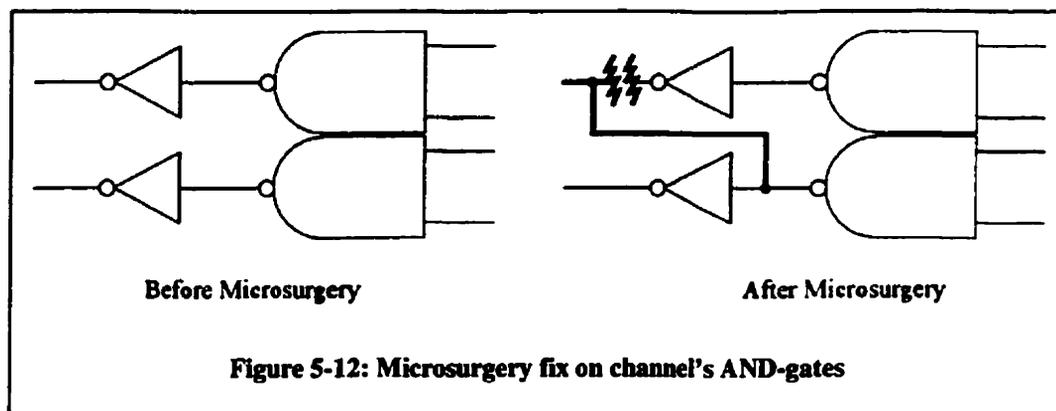


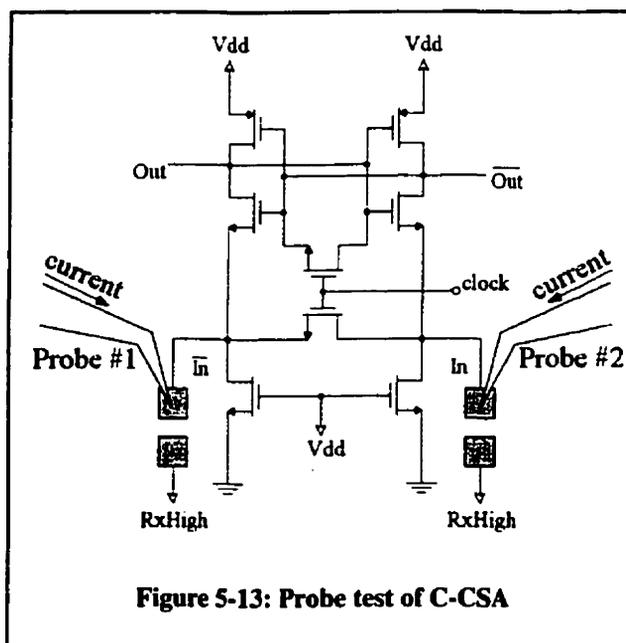
Figure 5-11: Digital error in P3A chip

secondary D-FFs with the control of the output concentrator multiplexers by using the complement of the D-FF control to control the multiplexer. There were a total of 14 procedures done to each chip, and involved using a focused argon-ion beam to pierce through the SiN_x and SiO₂ dielectric layers and then the metal trace line. A tungsten vapor was then passed into the argon-ion beam to build a connection to the other metal

trace line [figure 5-12]. The microsurgery was completely successful on two chips and was partially successful on one other. These chips were then post-processed with the MQW devices.



The second design flaw was a far more detrimental problem than the logic error. The performance of the clocked-charge sense amplifier (C-CSA) optical receiver circuitry within each smart pixel was very poor. During the preliminary all-electrical testing of the C-CSA, input currents were provided to both the optoelectronic contact points of the C-CSA using two flexible cat-whisker microprobes [figure 5-13]. Each microprobe was attached to an output on an HP4145B semiconductor parameter analyzer. The HP parameter analyzer could source and sink very small currents with a very high accuracy. Although simulations indicated that the C-CSA could operate with as little as 2- μ A and 4- μ A differential photocurrents [ref 7], the minimum experimental currents that were required to provide a stable logical output were much higher. The minimum currents were approximately 75- μ A and 25- μ A for a logical high, and 25- μ A and 75- μ A for a logical low. This implied that a pair of beams with 150- μ W and 50-



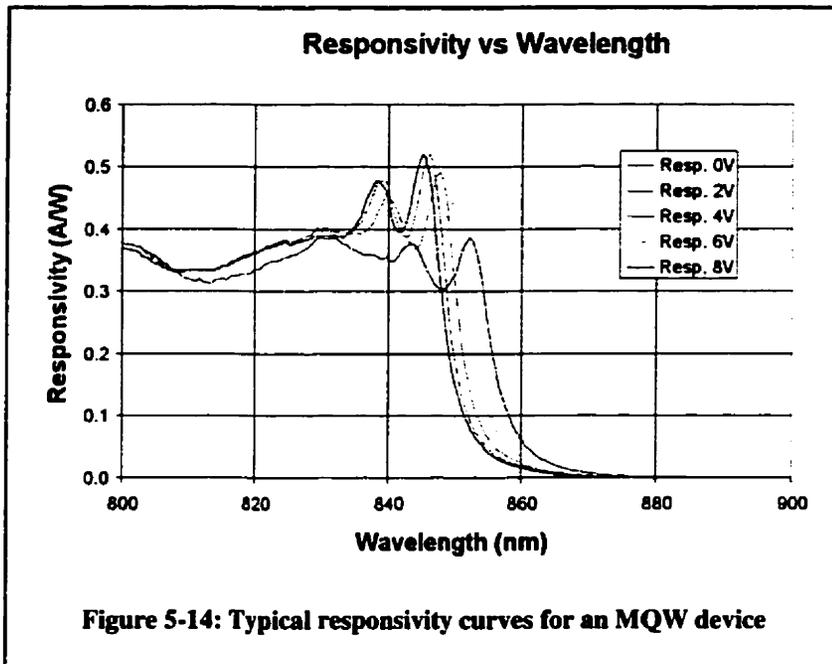
μW of optical power would be required assuming a 0.5A/W responsivity from the MQW diodes. Unfortunately, the Phase-III optical design could provide a differential power at the detector plane of approximately $40\text{-}\mu\text{W}$ and $20\text{-}\mu\text{W}$ given a 30-% / 60-% reflectivity from the modulator MQW diodes.

When the optoelectronic version of the P3A chip was tested, a very simple optical test rig with two complementary operated fiber-pigtailed VCSEL boards from Finisar™ were used. The highest speed attainable was approximately 37.5-Mbps with a clock rate of 75-MHz with optical powers in each beams of $0\text{-}\mu\text{W}$ and $130\text{-}\mu\text{W}$. This result was in close agreement with the preliminary all-electrical testing of the C-CSA, but the required optical power would be too high to be possible in the Phase-III optical system. In addition to this poor performance, it was found that the voltage bias for the detector MQW diodes, called RxHigh, did not affect the performance of the C-CSA. The RxHigh bias line could be completely disconnected from the power supply and it had no effect on the bit pattern. There seemed to be enough residual charge in the chip and the unconnected bias line to allow the C-CSA to latch on the input photocurrents. However, if the bias supply line RxHigh was increased above +6.5-V, the C-CSA would stop functioning altogether. Another disturbing feature of the C-CSA performance was that only one of the two modulated optical beams was required to change the state of the C-CSA. For example, if the C-CSA naturally produced a logical low value when no light was present, then only the detector MQW diode that was responsible for toggling the C-CSA in the opposite direction (logical high) required a modulated optical beam. Unfortunately, it was impossible to predict which diode in the pair was the useful diode because tendency to rail to either a 5-V or 0-V output appeared to be random across the chip.

The C-CSA's poor performance could not be analyzed in-depth because there was no way to experimentally examine the performance of a single C-CSA without the overhead of the logic of the entire smart pixel array. However, some possible reasons for the poor results can be outlined. Although the immediate loads on each C-CSA were identical in size, one of the complementary signals was inevitably disregarded because the CMOS circuitry required only a single output from the C-CSA. This could have skewed the output loads resulting in an asymmetry. Another reason for its poor performance

could have been differences between the resistance and capacitance of the MQW diode pairs. This could slightly skew the input impedance of C-CSA. Finally, the CMOS process fabrication may not have had a sufficiently tight tolerance on doping levels across the chip because it was not intended for highly sensitive positive-feedback elements like C-CSAs.

The final design flaw was once again the absolute reflectivity of the modulator MQW diodes. A detailed analysis of the MQW diode responsivity and reflectivity was performed on many devices over several chips. The best reflectivity obtained was $R_{low} = 7\%$ and $R_{high} = 15\%$ under optimal biasing and temperature conditions. Although,



measurements of the responsivity correlated well with the theoretical curves [figure 5-14] that indicate the MQW layer structure was intact. Unfortunately, the absolute reflectivities of the MQW devices were far from their designed

values. Further experimentation with the MQW diodes revealed the possibility of a poorly reflective gold mirror under the MQW stack. An experiment using a piece of GaAs substrate (prior to the flip-chip process) that had been polished and A/R coated on the backside was used to determine the approximate reflectivity of the gold mirror. An experiment was set up to measure the reflected and transmitted light through the sample when; i) the light passed through the bulk region of the sample, and ii) when light struck the gold mirror and reflected back through the bulk material.

The experiment used light at 904-nm so that the GaAs substrate would not absorb the light as it traveled through the substrate [figure 5-15]. The first set of equations below relate the transmission of the light through 500- μm of GaAs substrate and the reflectivity of the GaAs/SiN_x/Air interface. It was assumed that the A/R coated side contributed approximately 1-% of the reflected light and that the absorption of the GaAs was linear at the wavelength used (and not exponential).

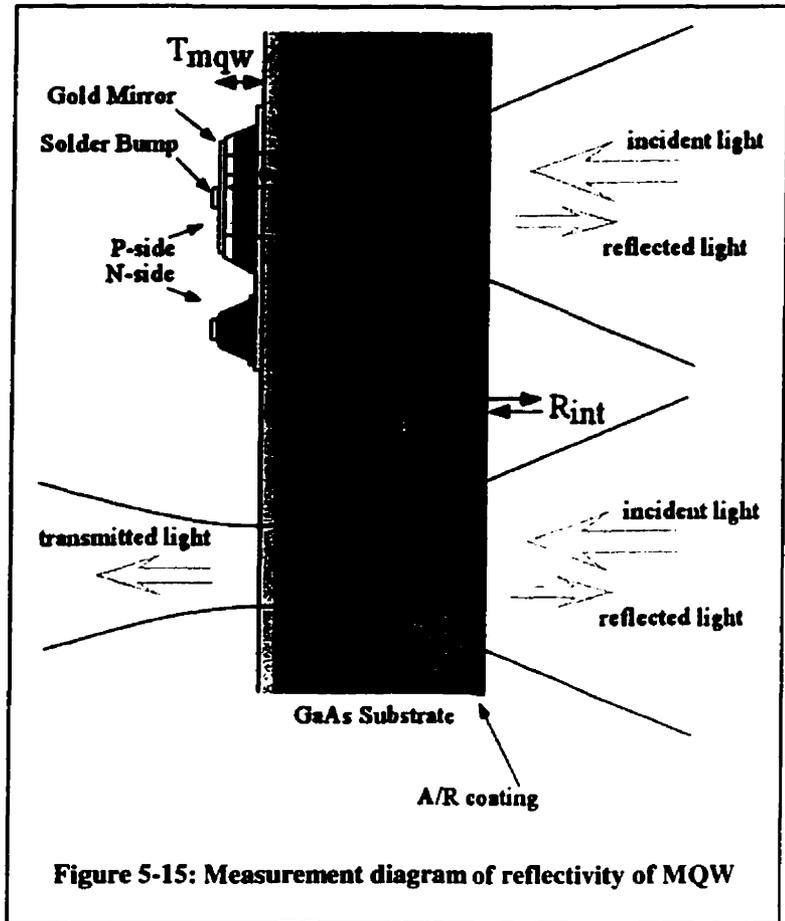


Figure 5-15: Measurement diagram of reflectivity of MQW

Eqn. 5-1: $0.029 = (0.99)(T_{SUB})(R_{int})(T_{SUB})$

Eqn. 5-2: $0.809 = (0.99)(T_{SUB})(1 - R_{int})$

This yielded an $R_{int} = 4.0\%$ and a $T_{SUB} = 85.1\%$. The next equation allowed an estimate of the gold mirror reflectivity to be made.

Eqn. 5-3: $0.153 = (0.99)(T_{SUB})(T_{mqw})(R_{mirror})(T_{mqw})(T_{SUB})$

This leads to a total device reflectivity of: $(T_{mqw})^2(R_{mirror}) = 0.213$. The mirror was therefore approximately 21.3-% reflective. If a 30-% to 80-% reflection was expected assuming a 100-% mirror, then using a 21.3-% reflective mirror would have produced a

6.4-% to 17-% reflection. This data correlates well with the experimental reflectivity measurements of the MQW modulator under optimal conditions.

The following photograph of an MQW diode cluster illuminated under an 860-nm LED light source shows regions of high reflectivity in the center of a few MQW diodes within a cluster [figure 5-16]. It is possible that during the flip-chip process, the unprotected area outside the immediate region of the solder-bump point was contaminated with the lead solder. Once this contamination had entered the GaAs, it spread over the central region on some of the MQW diodes creating completely dark MQW diodes.

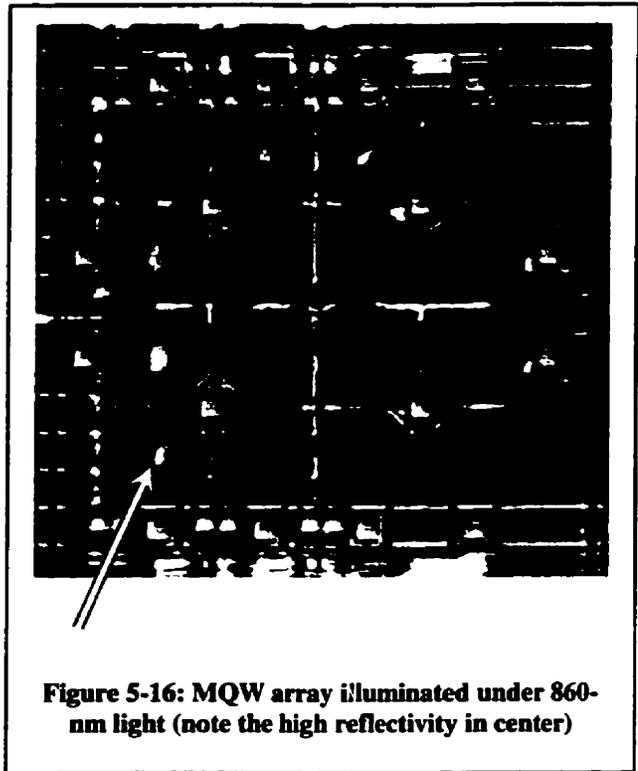


Figure 5-16: MQW array illuminated under 860-nm light (note the high reflectivity in center)

Due to a collection of errors in the design of the P3A chip as well as the optoelectronics, the P3A chip was never integrated into the Phase-III optical system. The performance of the chip as a whole would not have produced results of any real significance.

5.3.4) The P3B Chip

The P3B chip was first tested without the flip-chipped optoelectronic devices using the DB-Board test-rig and the cat-whisker microprobes under a microscope. All aspects of the digital design were verified as described in the previous chapter (see Chapter 4 – VLSI Optoelectronics). The next set of tests involved a three-beam optical test rig (OTR) that could interrogate the post-processed optoelectronic smart pixel array.

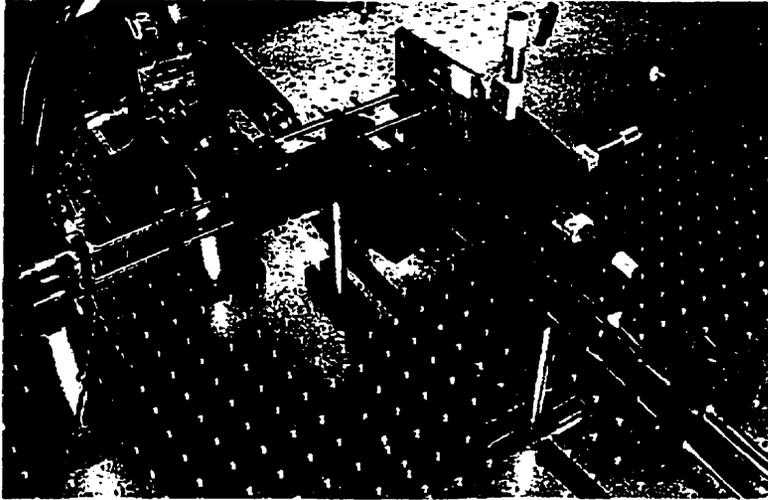


Figure 5-17: Photo of optical test rig

The OTR is shown here [figure 5-17] [figure 5-18], and allowed for 2 complementary modulated optical beams from two separate VCSEL sources to be relayed to the dual-rail optical receivers on the chip

(the DUT plane). The VCSELs were stimulated using complementary digital signals such that when one had a low output power, the other had a high output power. Output power from a VCSEL could range from 2- μ W to 90- μ W at the plane of the P3B chip, where typical pairs of optical powers were 60- μ W/20- μ W. The OTR also had a third optical beam that was used to interrogate the modulator MQW diodes. The third beam was a constant power read-out optical beam from a single-mode fiber. The source of the cw light was an SDL tunable diode laser that had been fiber coupled into the single-mode fiber. The output power of the constant beam could be varied from 0 to 4-mW. The constant beam was directed at a modulator MQW diode and the reflected power was redirected to an ANTEL high-speed avalanche photo detector. Only one constant beam was required for the

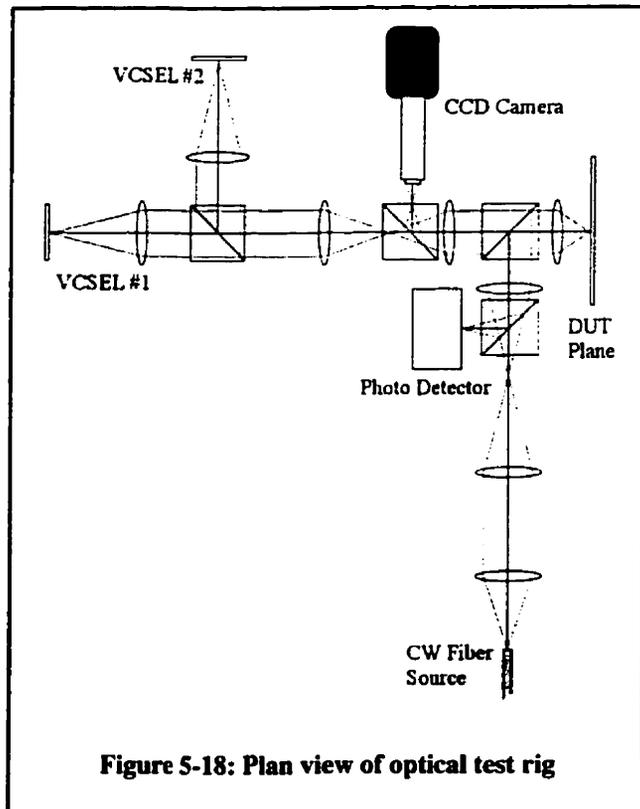


Figure 5-18: Plan view of optical test rig

detection of optical data from the MQW modulator diodes since the second complementary beam was redundant. This optical test rig was used to verify the high-speed optical operation of all three states of the smart pixel. Sending electrical data to a smart pixel and detecting the reflected light from the constant read-out beam with the photo detector tested the inject-state. The extract-state was tested using the complementary VCSEL beams on the detector MQW diodes and routing the signal off-chip to an oscilloscope.

The Inject-State:

By selecting the appropriate channel and smart pixel, electrical data was converted into a modulated light signal by the modulator MQW diodes. A single modulator MQW diode in the totem-pole would modulate the intensity of the incident constant 852-nm light and the optical test rig would redirect the reflected light onto an external photo detector. The electrical signal was generated on the Phase-III-B exercise board from a counter within the Altera FLEX10K programmable chip and an external clock source. The generated data was then routed along the Flex-PCB to the P3B chip [figure 5-4].

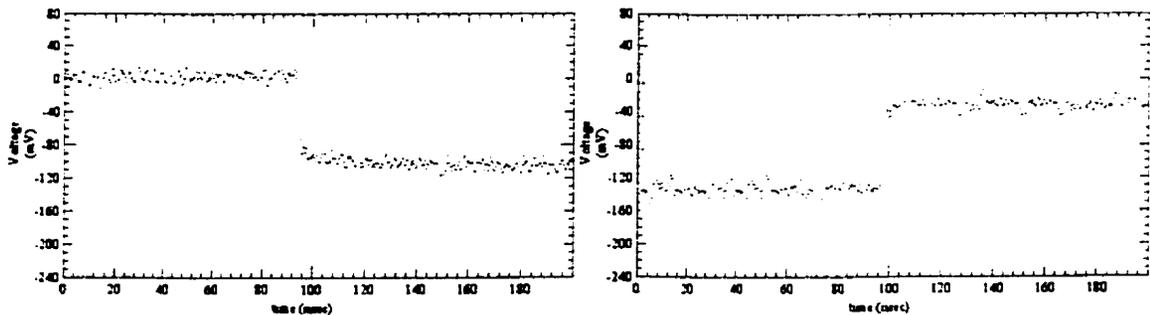


Figure 5-19: Measurement of fall-time and rise-time of electrical to optical modulation

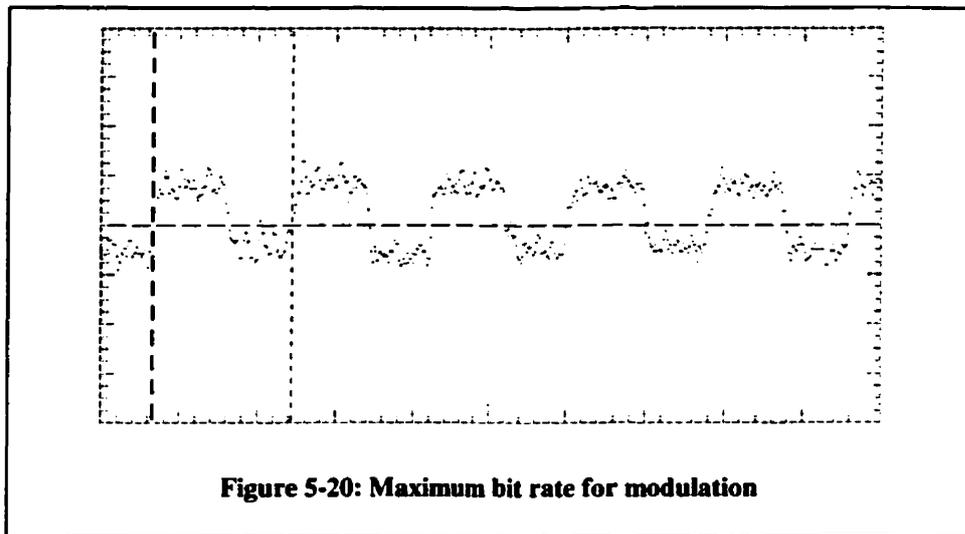
Using the ANTEL Avalanche Photo Detector, the measurement of the fall-time (rise-time) was determined. The observed rise-time was 1.65-ns and the observed fall-time was 1.8-ns as measured by the detector [figure 5-19]. By deconvolving the speed of the detector, the actual fall-time can be calculated by eliminating the bandwidth limit of

the digitizing scope and the photodetector where each is modeled by a first-order low-pass filter. The deconvolution is given by: $t_r^2 = t_d^2 - (t_p^2 + t_o^2)$.

Where t_r is the actual rise time, t_d is the observed rise time, t_p is the photodetector rise time (which is approximately 210-ps), and t_o is the digitizing oscilloscope rise time (which is approximately 7-ps). Given these times, the actual rise-time of the modulators was 1.637-ns and the actual fall-time was 1.788-ns.

The voltage scale on these graphs is somewhat unimportant, the voltage amplitude of each was approximately 15-mV. This corresponded to a power difference of about 4- μ W. Absolute power measurements of 8- μ W and 12- μ W at the ANTEL and approximately 32- μ W and 48- μ W reflected from the surface of the modulator MQW diodes.

The maximum bit rate of the modulator MQW diodes with the 1.5- μ m CMOS drivers was approximately 56-Mbps [figure 5-20]. This was a partially qualitative measurement, and based on the stability of modulated optical data waveform. This value



is very consistent with the quote speed of typical the 1.5- μ m CMOS Mitel Fabrication which quotes roughly a 50-MHz maximum clock rate.

The Extract-State:

Optical data was generated using an HP 80000 Data Generator connected to two VCSEL boards custom designed by Mr. Michael Ayliffe. The optical test-rig, discussed above, allowed a pair of complementary intensity light beams to be directed at the dual-rail MQW detector diodes within a smart pixel [figure 5-21]. The optical data was converted into electrical data by the optical receivers within the smart pixel



Figure 5-21: Two spots on MQW detectors

and routed through the Flex-PCB and the Mother-Board to a digitizing oscilloscope. The maximum bit rates were limited to the speed at which the P3B chip could operate. This was approximately 50-MHz due to the rather large transistor line-width of 1.5- μm . The bandwidth of the Flex-PCB and the Mother-Board combination did not limit the speed of

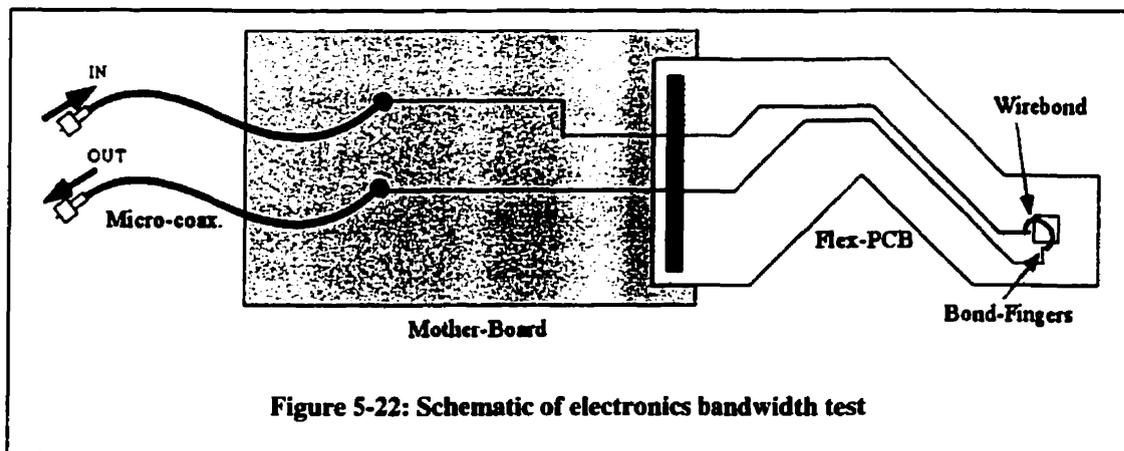


Figure 5-22: Schematic of electronics bandwidth test

the modulated data. An experiment verified the bandwidth by using a pair of designated data lines to and from the bond-fingers near the chip. A single wire-bond between these two bond-fingers allowed data to be passed into the Mother-Board and back-out again to verify the speed of the electrical connections [figure 5-22]. The eye-diagram in [figure 5-23] shows a very high bit rate. The bit period was 1-ns and the PRBS code length was $2^{64}-1$.

To measure the characteristics of the optical receiver in its most-sensitive state, the optical receiver was used in its open-loop mode [figure 5-24]. The feedback pMOS

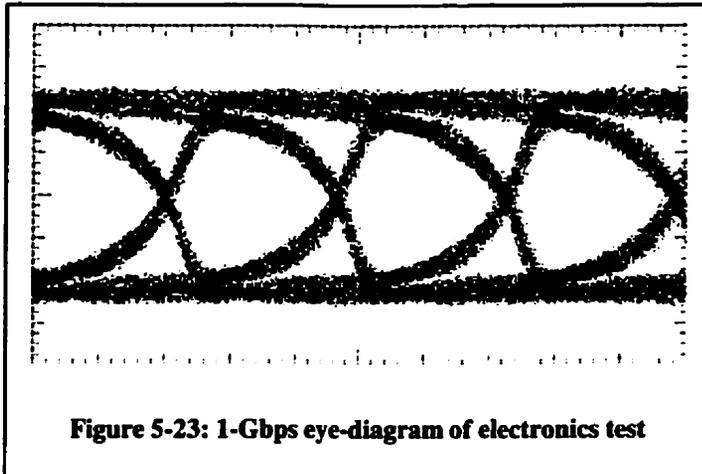


Figure 5-23: 1-Gbps eye-diagram of electronics test

resistor was open-circuited by keeping the gate of the pMOS transistor at 5-volts. This data was obtained by measuring the performance of the individual optical receivers that had been placed on the P3B chip using their own set of directly accessible output bond pads.

This allowed for a very sensitive optical receiver, however, the bit-rate was severely limited due to the slew-rate of the open-loop amplifier and the lack of an automatic threshold mechanism. The diagram in [figure 5-25] is representative of the slew-rate limit

with open-loop amplifiers. If a single logical zero bit follows a train of logical one bits and proceeds another train of logical one bits, it can be lost due to the low rate of change of the voltage per time of the amplifier. Although the maximum data rate for the open-loop amplifier was very low, less than 1-MHz, the sensitivity

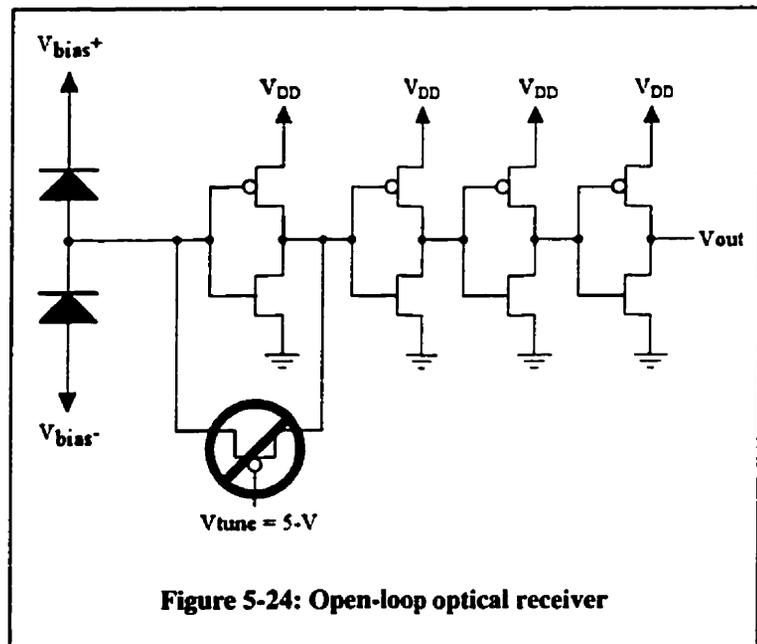


Figure 5-24: Open-loop optical receiver

of the amplifier was extremely good. The open-loop optical receiver on the P3B chip could sense a differential optical power down to approximately 2- μ W and 4- μ W [figure 5-26]. The rise-time for the open-loop amplifier was approximately 1.43- μ sec and the slew-rate was 8.2-V/ μ sec. The high RC-constant, due to the input capacitance of the

inverter-amplifier and the parallel combination of the large capacitance and resistance of the MQW, limited the data-rate of the signal. However, the gain of the amplifier was very high (roughly 25-V/V), and therefore the detector was exceptionally sensitive at low frequencies.

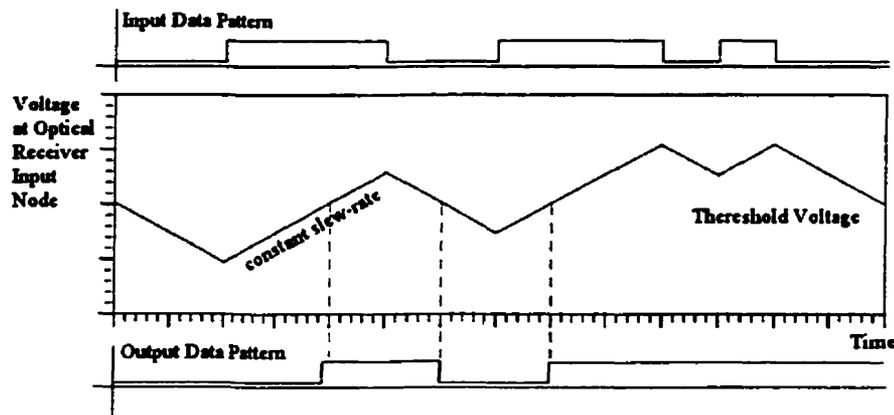


Figure 5-25: Effect of open-loop amplifier slew-rate

To obtain the highest possible data rate, the optical receivers were placed in their normal mode of operation by applying zero volts to the gate of the feedback pMOS transistors. Optical receivers within smart pixels were tested using optical data from the VCSEL boards. The optical data was incident on the optical receiver of a smart pixel

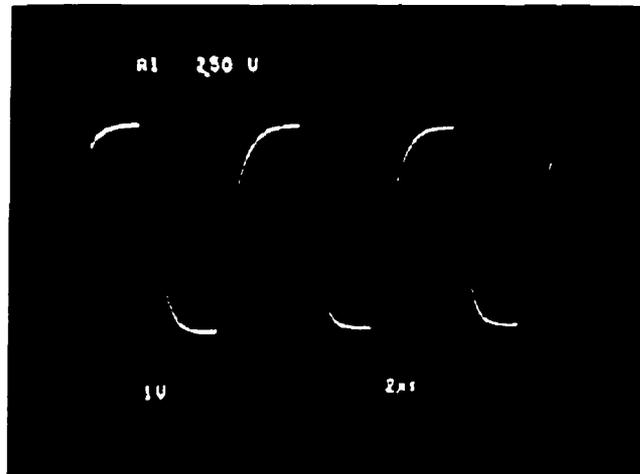


Figure 5-26: Open-loop data pattern at 1-Mbps and 5-V swing

and converted into electrical form. The data was then routed off the chip, along the Flex-PCB, and out the Mother-Board to a high-speed digitizing oscilloscope. Several sets of data were obtained from these.

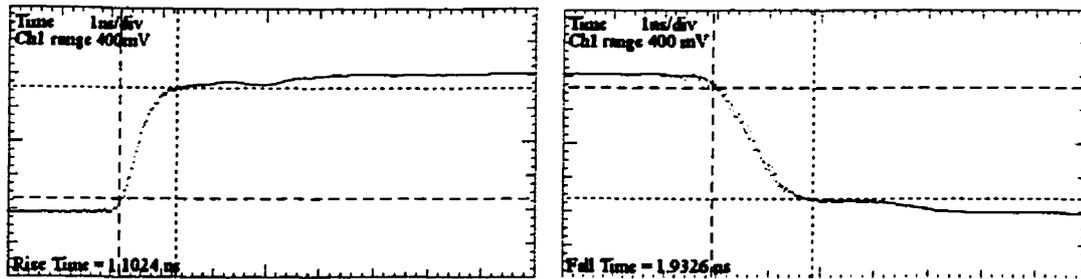


Figure 5-27: Rise and fall times for smart pixel extraction state

The minimum differential optical power required to switch the TZA within the smart pixel was approximately 24- μ W and 44- μ W. The rise-time of the optical-to-electrical signal was 1.10-nsec and the fall-time was 1.93-nsec [figure 5-27]. A square-wave pattern as well as a 16-bit pattern were also obtained [figure 5-28] [figure 5-29] and have data rates comparable to the maximum bit-rate for the technology. Eye-diagrams were also obtained and showed relatively clean, open eyes up to 60-MHz [figure 5-30] [figure 5-31]. However, since the dual-rail encoded optical beams of light were obtained from two

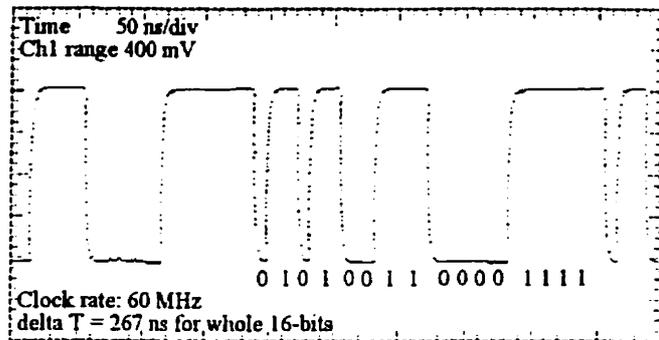


Figure 5-28: 16-bit pattern for smart pixel extract state

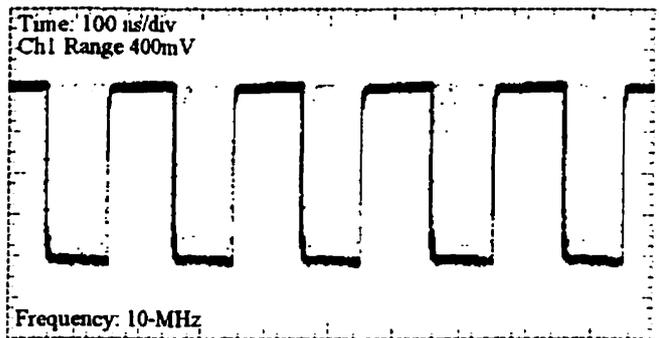


Figure 5-29: Square wave for smart pixel extract state

independent VCSEL boards, and each board require two complementary signals, the PRBS code for all four channels from the data generator were manually entered. The eye-

diagrams appear slightly deformed because the bit-patterns did not represent a true PRBS code. A second difficulty was the synchronization of all these bit patterns at high

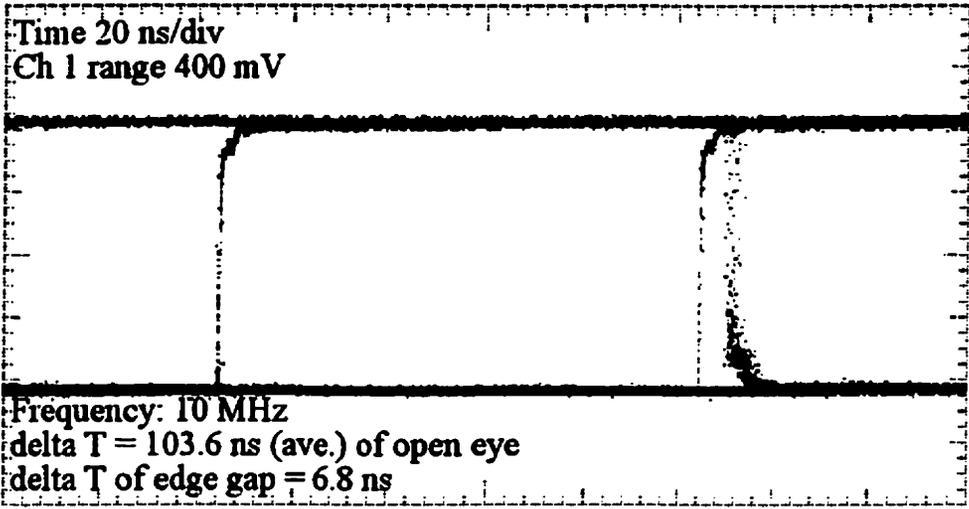


Figure 5-30: Lower-speed eye diagram for smart pixel extract state

frequencies and thus the absolute power in the optical beams was not perfectly

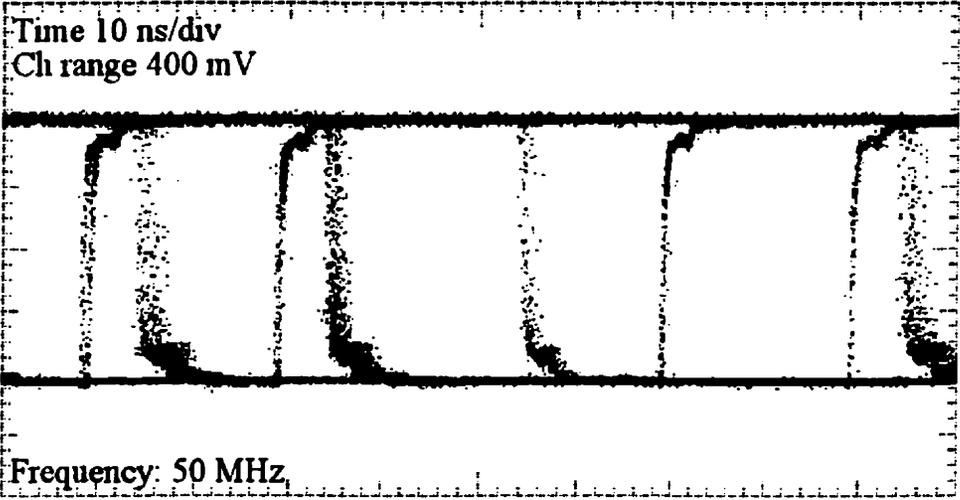


Figure 5-31: Higher-Speed eye diagram for smart pixel extract state

complementary.

5.4) Conclusion

In this chapter, evidence of fully functioning VLSI-OE chips was presented. However, in addition to data collected on bit-rate, rise time, and optical power requirements for optical receiver circuits, several testing strategies and testing apparatus were highlighted.

Test set-ups that involved the testing of VLSI-OE chips prior to the integration of optoelectronic devices were constructed. The non-destructive testing of VLSI-OE chip without optoelectronics devices was essential, since the process of integrating optoelectronics with the chip had to be carried out on fully functioning chips – a cost as well as a time constraint. Electrical tests of the receiver and transmitter circuits that would eventually be connected to the optoelectronic devices were essential in the verification of the design. This allowed the step-by-step testing to be carried out on all subcircuits on the chips. This testing was carried out using specialized microscope based platforms using microprobe tips to excite and detect the voltages at the optoelectronic contact points on the chips.

In addition to purely electronic testing, the VLSI-OE chips were also interrogated using optical input and output data along with electrical input and output. These signals must be properly timed and are usually coordinated with a global system clock. The generation of optical data must be done using external light sources, such as packaged VCSELs. Appropriate photodetectors, external to the chip, must also be used to detect output optical signals. Therefore, a properly calibrated custom optical test set-up was required. The optical test-rig allowed high-speed electrical signals to and from the chip. The completely functioning optoelectronic chips were then integrated into the optical systems with a high confidence that they indeed worked properly.

The testing of the four VLSI-OE chips helped provide a methodology and proficiency with the design and analysis of free-space optoelectronic microchips. The four designs provided information on a design's sensitivity to floating inputs (as in the Workshop-Chip), on the performance of transimpedance and current sense amplifiers (as in the P3A and P3B chips), and on the sensitivity to misalignment when using small optoelectronic devices (as in the Beta-Chip).



5.5) References

- [1] A. Okuno, K. Nagai, K. Ikeda, T. Tsukazaki, N. Oyama, K. Nakahira, T. Hashimoto, "New packaging of chip-on-board by unique printing method", Proceedings of the 41st Electronic Components and Technology Conference, 1991, pp. 843 -847
- [2] D.N. Kabal, G.C. Boisset, D.R. Rolston, D.V. Plant, "Packaging of two-dimensional smart pixel arrays", IEEE/LEOS 1996 Summer Topical Meetings, 1996, pp. 53 -54
- [3] M.H. Ayliffe, D.R. Rolston, E.L. Chuah, E. Bernier, F.S.J. Michael, D. Kabal, A.G. Kirk, D.V. Plant, "Packaging of an optoelectronic-VLSI chip supporting a 32x32 array of surface-active devices", Optics in Computing 2000, SPIE Vol. 4089, Quebec City, Canada, 2000, pp. 508-518.
- [4] D.V. Plant, B. Robertson, H.S. Hinton, W.M. Robertson, "An optical backplane demonstrator system based on FET-SEED smart pixel arrays and diffractive lenslet arrays", IEEE Photonics Technology Letters, Vol. 7, No. 9, Sept. 1995, pp. 1057-1059.
- [5] D.R. Rolston, D.V. Plant, T.H. Szymanski, H.S. Hinton, W.S. Hsiao, M.H. Ayliffe, D. Kabal, M.B. Venditti, P. Desai, A.V. Krishnamoorthy, K.W. Goossen, J.A. Walker, B. Tseng, S.P. Hui, J.E. Cunningham, W.Y. Jan, "A hybrid-SEED smart pixel array for a four-stage intelligent optical backplane demonstrator", IEEE Journal of Selected Topics in Quantum Electronics, Vol. 2, No. 1, Apr. 1996, pp. 97-105.
- [6] Texas Instruments: Digital Logic Technology Families – HC/HCT High-Speed CMOS Logic, <http://www.ti.com/sc/docs/products/logic/families/hct.htm> , 1999
- [7] Alain Z. Shang, Transceiver Arrays for optically interconnected electronic systems, Ph.D. Thesis. McGill University, Montreal, Canada, 1997.

Chapter 6: Synchronization

6.1) Introduction

Many high-speed digital switching systems and especially massively parallel processing systems in use today require some form of complicated clock synchronization [ref 1]. A reference frequency within a computer is used to sequence events such that data and instructions are ordered in a particular fashion. Typically, a clock signal is distributed to all the parts of the system to ensure that all subsystems execute their instructions in the correct order. There are many forms of ynchronized systems, ranging from microprocessor chips to printed circuit boards (PCBs) and backplanes, to long-distance telecommunication systems such as optical fibers and satellites. Each system requires different synchronization methods in order for their subsystems to interact properly; where each may have their own definition of synchronization.

In this chapter, a new method for synchronizing systems, called the *distributed synchronous clock (DSC)*, will be presented. The definition of synchronous operation for the DSC will be the following:

- 1) *A periodic pulse train is distributed, in some manner, to every node in the system.*
- 2) *Every node in the system receives a rising/falling edge of a pulse (in-phase) at precisely the same moment in time regardless of the relative distances among the nodes.*
- 3) *Every node in the system receives the same frequency periodic pulse train.*

The impetuous behind the development of this technique originated from VLSI-OE chip design and testing. In the architecture outlined earlier (see Chapter 2 – Architecture) the smart pixel array required multiple in-phase clocks that had to be delivered to many physically distance points without any skew. In order to provide a synchronous clock to each node in the system without constructing a secondary clock distribution network, a method to distribute the clock using the existing closed-loop optical ring was sought. This excluded the use of external synchronization techniques,

such as H-trees, and required a method by which the distribution network itself would aid in its own synchronization.

This chapter will begin with a review of several standard synchronization techniques used in present day computing systems and will discuss the basic clock requirements for most computing systems. Several existing clocking methods will then be described which offer alternatives to typical clocking structures. The fundamental concepts of the distributed synchronous clock will then be discussed. This development will begin with a review of the digital ring oscillator. Starting with the digital ring oscillator as the initial design, a series of simple additions or modifications will be made to overcome problems associated with previous ideas in the progression. The final circuit will then be presented along with HSpice simulations and an example based proof of concept for the clock mechanism. A treatment of a simple digital phase-locked loop (DPLL) will be given in the appendix at the end of this chapter. The DPLL is an integral part of the distributed synchronous clock and should be explored.

6.2) Standard synchronization techniques

6.2.1) On-chip synchronization

The DEC-alpha and the Pentium Series of microprocessors are examples of systems that use on-chip clock generation and distribution to synchronize their subsystems [ref 2,3]. A clock multiplier circuit, internal to the microprocessor, is responsible for the generation of high-speed internal clocks from lower speed external clocks. A frequency altering digital

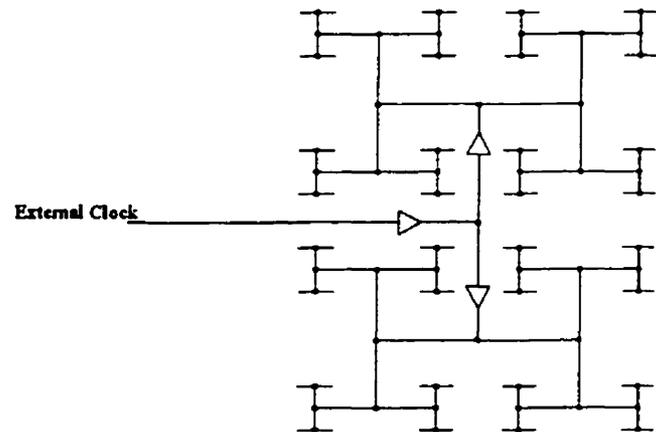
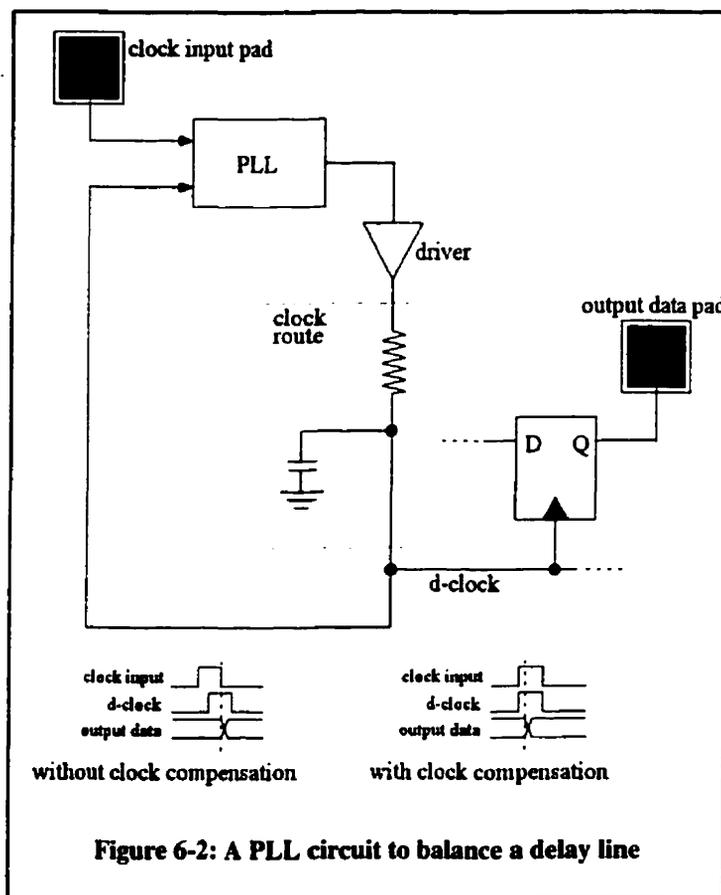


Figure 6-1: H-tree clock distribution

phase-locked loop (DPLL) is the typical way a high frequency clock is generated from a low frequency clock. The low frequency clock is attached to an input bond pad of a

microprocessor chip from an external source. This clock frequency is then increased on the chip using a frequency multiplying DPLL that is subsequently distributed to each of the subsystems within the chip. The frequency altering DPLL is described in some detail in the appendix at the end of this chapter.

The technique used to distribute high frequency clocks within these chips with very small amounts of skew among all the end-points of the distribution network is typically a length equalized binary or H-tree structure [figure 6-1]. This type of layout equalizes the delay of the clock pulses to every point on the chip. However, this requires that all clock buffers have identical delay characteristics and that all paths are equal [ref 4]. Another method used to overcome clock skew involves the partitioning of systems on the chip so that each subsystem has its own synchronized clock. Each subsystem must



then interact with other subsystems using a protocol, which typically involves memory buffers and control signals. A third method to synchronize clock pulses employs a PLL to measure and adjust the phase of a clock in order to synchronize different subsystems [ref 5] [figure 6-2]. Unfortunately, the use of a PLL is limited because the PLL circuit can consume a large amount of area on a chip.

To highlight the problems associated with clock skew within a microprocessor, the 1999 Pentium III processor will be used as an example of a system that would suffer from significant skew if some of the previous clocking structures were not employed. The Pentium III has an internal clock rate of 600-

MHz [ref 6], which means that less than a 2-nsec period square-wave pulse train must be distributed to each register within its processing pipelines. Assuming that a maximum allowable skew of 5% is tolerable among the clock edges arriving at each register, the clock pulses must arrive in less than 100-psec with respect to one another. If the propagation velocity of a microstrip line on the chip is roughly 20-cm/nsec, and the greatest distance a pulse must travel is approximately 2.5-cm. The maximum possible skew is 125-psec. This value is comparable to the 5% maximum skew of 100-psec, and hence it is very possible to violate the minimum skew condition. This example serves to illustrate an extreme case of skew sensitivity. Although these issues can still be addressed using standard for chip-level skew, it is extremely difficult to apply chip-level solutions to larger systems while maintaining the high clock rates.

Synchronization among several chips is also a very critical issue. Many high-speed CPU and DRAM memory interconnections are packaged in the same container in order to keep all the subsystems requiring the same clock, as close as possible. The multi-chip module (MCM) is a packaging technique that has at least two interconnection levels and at least two chips, such as a memory and a CPU. An example of such a structure using a Pentium II

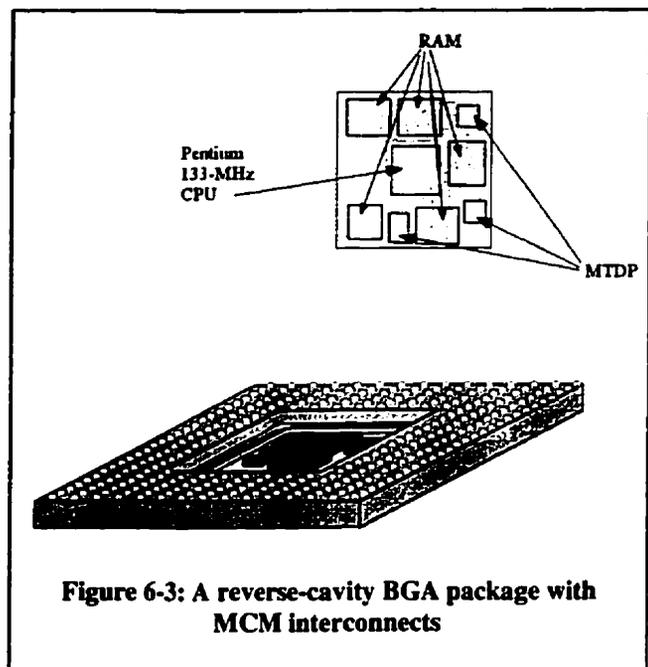


Figure 6-3: A reverse-cavity BGA package with MCM interconnects

processor and high-speed memory is shown below [ref 7] [figure 6-3]. The first interconnect layer is composed of the interconnect lines within the silicon chips, and the second interconnect layer is composed of trace lines made within the chip carrier. The trace lines within the package allow the chips to communicate with each other without any of their signals ever leaving the package. The number of external pins is reduced while the internal data and clock rates are increased. However, this introduces an increase

in complexity within the package and a decrease in the yield and internal diagnostic access.

A similar method to the MCM packaging technique, used to maintain speed and synchronization, is to place packaged chips beside each other in the form of a module. The module is a specially designed printed circuit board that uses isolated, well-designed microstrip bus lines and clock lines to communicate among the packages. The SUN Ultra-SPARC III processor [ref 8] is an example of a module that is capable of maintaining a data rate of about 155-Mbps per line between CPU and DRAM. The module is 10-cm on a side and has two connectors to the external system. Both connectors operate at speeds lower than 100-MHz. The PCI bus connector can only operate up to 66-MHz and is the main interface to the rest of the computer [figure 6-4].

The newest trend in system integration has been the introduction of the “system-on-a-chip” method. An example of system-on-a-chip integration is the IBM Core PowerPC 405 series [ref 9] and is one way to aid in the synchronization of multiple subsystems. The system-on-a-chip method is the next advance in IC integration and includes all the subsystems of the

MCM approach within the same silicon substrate [figure 6-5]. The system-on-a-chip method is attractive primarily because of the problems associated with the low yield of many chip packaging techniques such as the MCM [ref 10] and because present day transistor packing density makes this technique possible (Altera APEX chip set has on the order of 2.5-million gates) [ref 11]. One of

the key features of this technology is that it allows very high frequency synchronized clocks to be maintained among all the subsystems because the system is a monolithic chip.

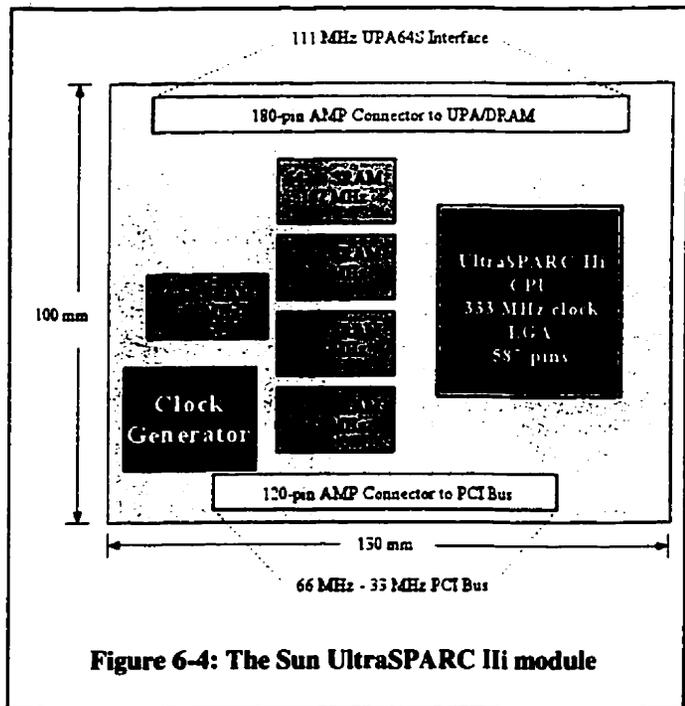


Figure 6-4: The Sun UltraSPARC III module

6.2.2) Board-to-board and computer synchronization

Although the challenges associated with high-speed chip design are numerous, the techniques for internal chip synchronization remain adequate for continued growth in this field for the time being. However, serious challenges to synchronization arise when several physically separated printed circuit boards are required to communicate together. The separation between points on two PCBs in a backplane interconnect can be up to one meter, whereas internal chip interconnects are typically a few millimeters. Thus, the synchronization techniques employed in chip level layouts are not well suited for PCB layouts. The main concerns with distributing high-speed data or clocking signals are the amount of power required and the integrity of the signals. Both the power and the signal integrity are greatly affected by mismatches in the impedance of the transmission lines. The electronic bus or backplane is a very common structure that is used to pass data among boards. A typical bus structure uses a hand-shaking protocol, many control lines,

and a bus arbiter to connect PCBs.

The VMEbus, the MULTIBUS, the FutureBUS+, and the PCI bus [ref 12,13,14] are all examples of typical structures used to connect PCBs together. These bus structures are usually master-slave implementations and require several layers of protocol in order to complete data transactions. A typical sequence of events in a bus transaction, modified from the VMEbus protocol, is given here [Table 6-1]:

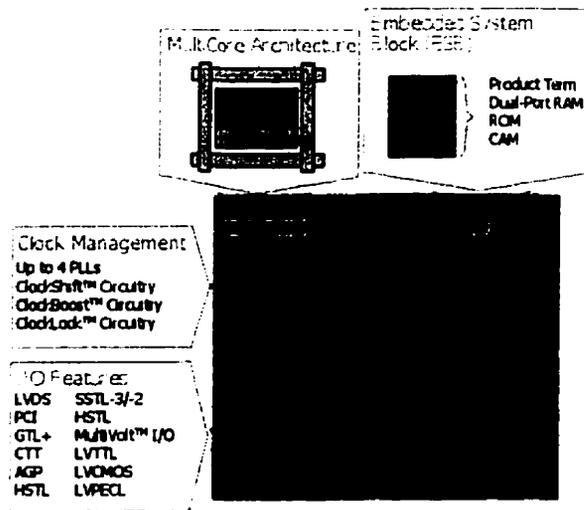


Figure 6-5: Photo of system on a chip

(<http://www.altera.com/html/products/apex2.html#arch>)
(as of March 3rd, 2000)

- The microprocessor of the Mth PCB initiates a memory access and outputs an address.

- The address decoder on the M^{th} PCB determines whether the address is for an on-board resource or for a bus resource. If the bus is required, it selects the requester.
- The bus requester of the M^{th} PCB requests the bus from the bus controller board.
- The bus control, some time later, grants the bus to the M^{th} PCB.
- The bus requester of the M^{th} PCB enables the interface logic to the bus allowing the microprocessor to broadcast the address onto the bus.
- The N^{th} PCB recognizes the request and connects to the bus.
- The microprocessor on the M^{th} PCB accesses the resources on the N^{th} PCB for some arbitrary duration.
- The N^{th} PCB requester releases the bus.

Table 6-1: Simplified Protocol of a bus transaction

The protocol described above requires a relatively large number of “extra” components to maintain an orderly flow of data between PCBs. Since there is no global set of perfectly synchronized clocks distributed to each board, additional components as well as a protocol must be used to regulate the transmission of data. This system of communication can be very robust, but there is a great deal of latency incurred. Any pair of PCBs that are waiting for a transaction to end must buffer their requests and hence spend some time idle while other PCBs are serviced. Although companies such as Nortel Networks have overcome some of the difficulties associated with transmission line effects, and have been able to construct backplanes that run in excess of 1-Gbps per trace line, absolute synchronous operation is still elusive. Most common electronic bus structures, such as the VMEbus and the PCI Bus support clock frequencies of less than 100-MHz and PCB synchronization is not even attempted [ref 15].

In a multiple processor system, such as the multiple board Cray T3E supercomputer [ref 16], the same clock distribution method used in microprocessor chips is also used to synchronize all the PCBs. A binary tree composed of single mode, dispersion compensated optical fibers from a single laser source is used to supply synchronous pulses to each board in the system. The fiber undergoes many one by two splits with pulse amplification until the signal can be distributed to all the PCBs in the system. The Cray T3E supercomputer was required to have less than 20-psec skew among all boards in a system with over several hundred processors.

6.2.3) Long-distance synchronization

In large, distributed networks such as long-distance telecommunications, the asynchronous transfer mode (ATM) protocol is used. ATM is only partially asynchronous and still requires clock recovery and memory buffers at each node in the network in order to sequence events. The asynchronous aspect of these switching systems means that nodes in the network are not synchronized together. Each node transmits data encoded with a clock signal at a nominal rate; the receiver node must lock to this frequency and interpret the incoming data stream. The internal processing of each ATM node is in fact very well synchronized. The data is encoded such that PLL clock-recovery methods can be used to extract a clock from the data signal. This clock is then used to sequence the incoming data streams [figure 6-6].

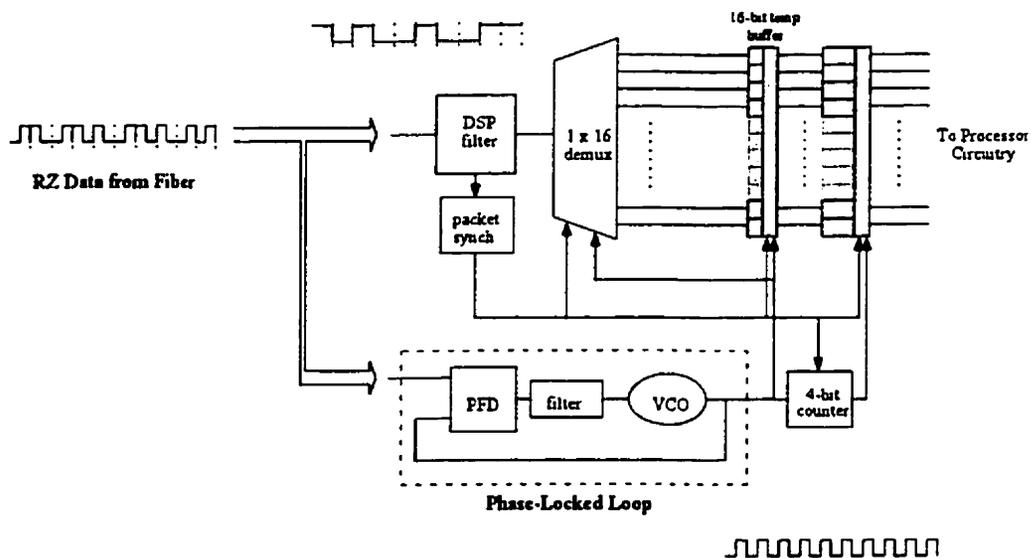
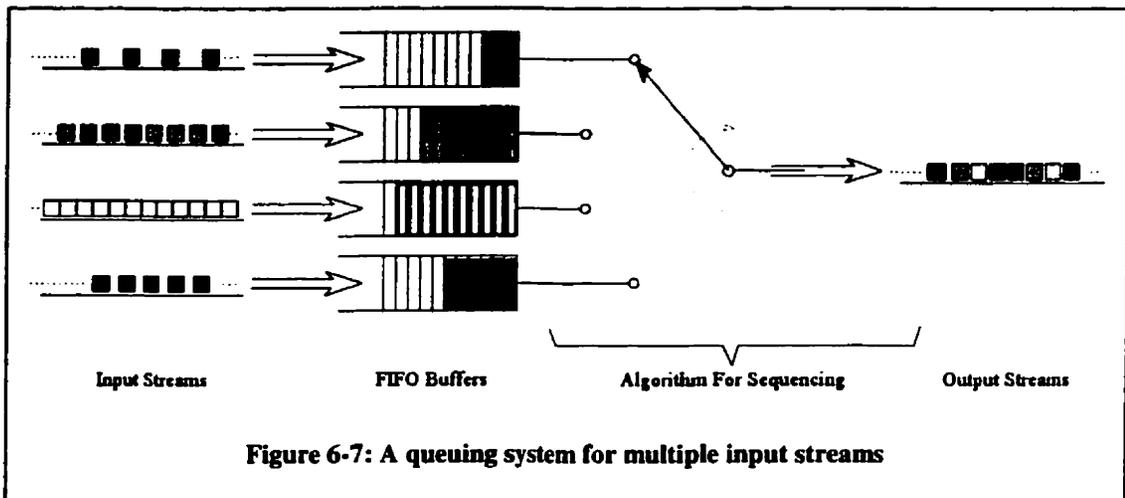
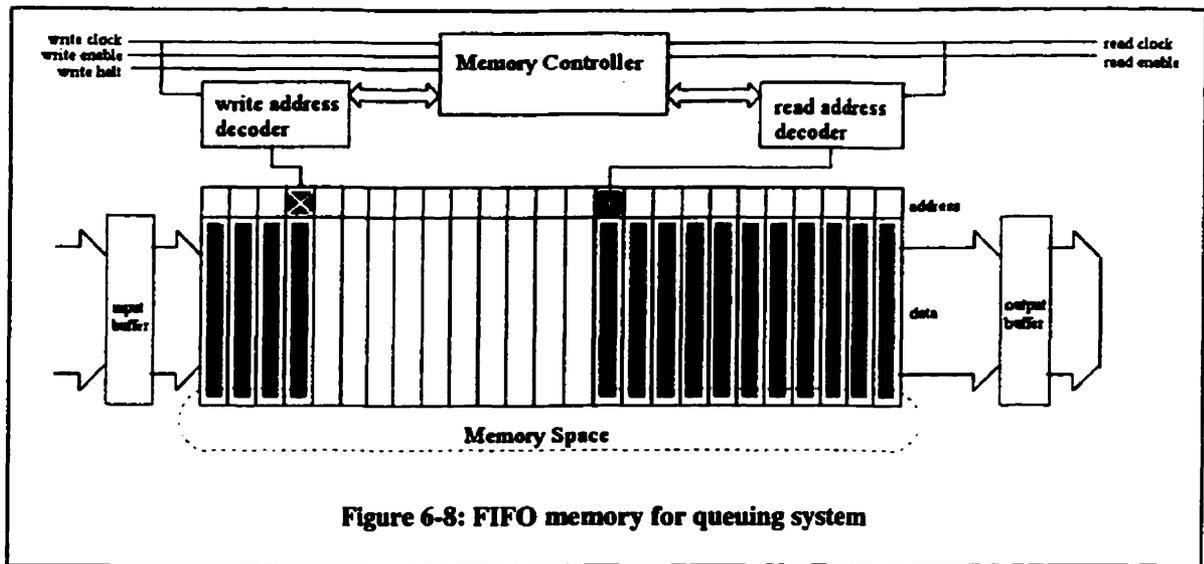


Figure 6-6: Concept of a clock-recovery circuit for optical fiber

Data buffers called circular First-In First-Out (FIFO) memory allow one subsystem, or node, to communicate with another without the need for absolute synchronization. The circular FIFO can match two slightly different frequency clocks by allowing the buffer to be filled using the incoming clock frequency and emptied with the out-going clock frequency [ref 17]. This is especially useful for systems passing asynchronous transfer mode (ATM) packets [ref 18]. Such a situation occurs when the clock rates of the input data streams to an ATM node are different from one another, and the output data stream uses a reference clock rate within the node [figure 6-7]. When incoming data is presented at the memory buffer, the write-state is enabled and the first clock is used to place data into the memory. When the data is read into the subsequent



system, the read-state is enabled and the second clock is used [figure 6-8]. However, this method requires a significant amount of overhead in circuit design to ensure that the buffer is never over-written, and requires status and halt signals for both subsystems.



6.3) Alternative clocking structures

The methods described in this section are alternatives to the more standard designs outlined in the previous section. Some of these methods are not widely used, but they offer unique perspectives on clock design. These examples are provided in order to compare some of their features with the distributed synchronous clock presented later in this chapter.

A clocking scheme which uses a network of inverting amplifiers called a cooperative ring oscillator (CRO) was prototyped by L. Hall et al. [ref 19] and is based on the three element ring oscillator [figure 6-9]. The three-element ring oscillator is an

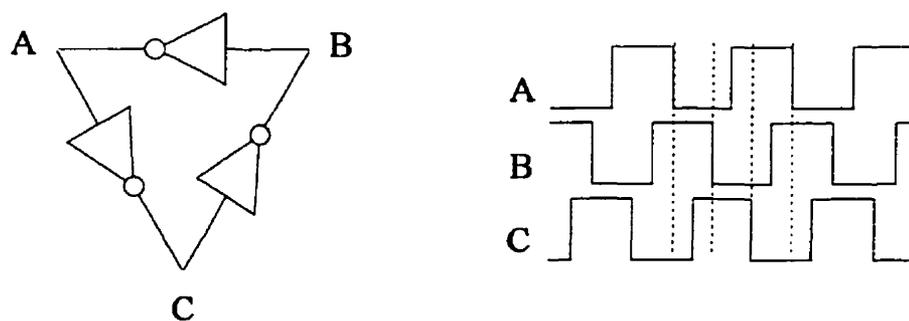


Figure 6-9: A three-element ring oscillator and waveforms

unstable circuit that spontaneously begins to oscillate during power-up. An odd number of 180° phase-shifting amplifiers (i.e.: an inverting amplifier such as the CMOS inverter) will cause an oscillation due to positive feedback.

By extending this principle in a two-dimensional fashion, each node in the circuit becomes a point in two other ring oscillators, such that each node has three inputs and three outputs [figure 6-10]. The entire network of inverters will begin to oscillate in the same way in order to satisfy the rise and fall of voltages at each node simultaneously. This structure can provide both the generation and the distribution of the clock signals and is ideally suited for a chip level distribution network because both the signal lines and the CMOS

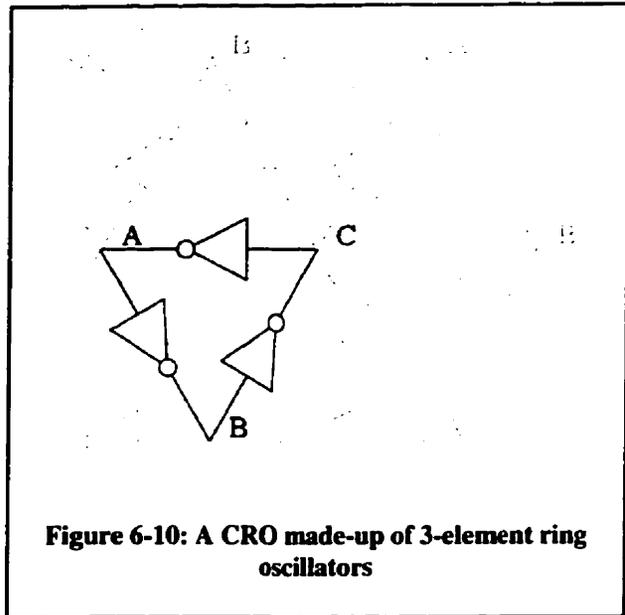


Figure 6-10: A CRO made-up of 3-element ring oscillators

amplifiers can be integrated onto the same substrate. However, this design is not suitable for large networks because of the finite propagation velocity of the signals. When the total distance of the network is less than a small percentage of the propagation velocity times the period of oscillation then the network will remain in-phase. However, when the distances become too large, the pulses that propagate through the network cannot all simultaneously reach all nodes. Some nodes can be rising while other corresponding nodes are still falling because the phase front has simply not arrived. Assuming that a phase mismatch of approximately 1% can be tolerated (this is roughly a 100-psec skew for a 100-MHz signal), then $d = 0.01v_pT$, where 'd' is the size of the network, ' v_p ' is the propagation velocity, and 'T' is the period of oscillation. Since a typical microstrip transmission line has a propagation velocity of about 20cm/nsec, the maximum distance that can be tolerated is about one centimeter. This is an upper limit on the total distance of the network because of the non-ideal transmission line effects of real interconnects on chips.

Another clocking strategy, proposed by W.D. Grover [ref 20], is based on the time of flight of a pulse in a single conductor. In this design, distance is not an issue as it was for the cooperative ring oscillator. The design is based on the transmission of pulses along a single conductive line such as a coaxial cable [figure 6-11].

At the beginning of the cable is a pulse generator that sends pulses into the line. The pulse travels down the conductor passing all the nodes in the system once. Once the

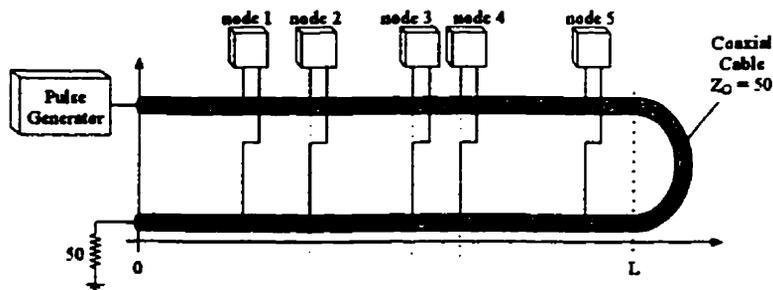


Figure 6-11: Grover patent based in time-of-flight delay halving

pulse has reached the mid-point of the line, it returns towards the source passing all the nodes again, but in the reverse order. The first node on the line is the first to receive the outgoing pulse and the last to receive the returning pulse. Using this strategy, any node in the path receives two pulses in time where the time between arrivals is different for each node. There is a longer interval between arrivals of the pulse for the near node, and shorter interval for the far node. However, when the mid-time between arrivals of a pulse at any node is calculated, it is found to be a constant and can be used as synchronous point in time for all nodes in the system [figure 6-12]. There are a few drawbacks to this

design; the first is the repetition rate of the pulse generation. The pulses can only be issued once per round trip otherwise they will interfere with one another if issued faster

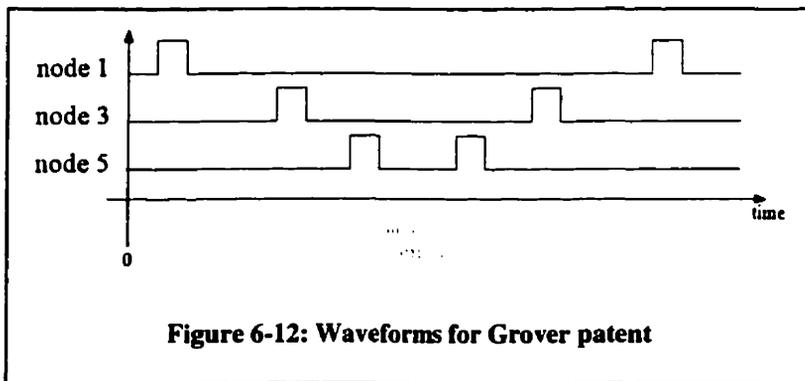


Figure 6-12: Waveforms for Grover patent

than the time of flight of the conductor. A second issue is that proper synchronization assumes that the transmission line properties remain constant along the entire path. This may be a reasonable assumption, however there are no mechanisms to determine if this is indeed true. It is possible for thermal variations or time degradation to change parts of the conductor's propagation velocity either statically or dynamically. This can also happen if the loading of the transmission line changes such as the case when a node is introduced or

removed from the system. Another drawback is the nature of the transmission medium itself. For highly integrated systems, a large coaxial conductor carrying the clock signal may not be congruent with modern integrated technology.

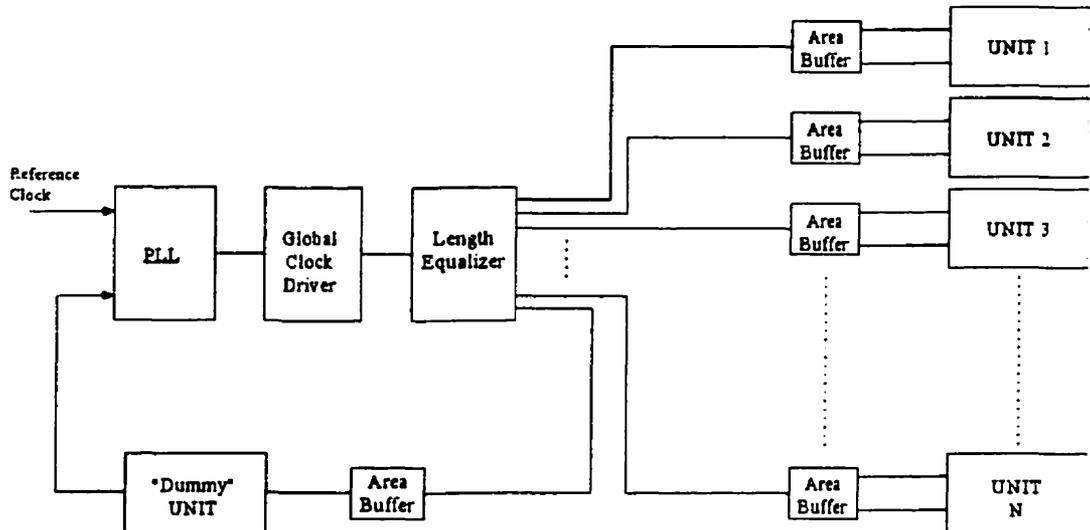


Figure 6-13: Clock equalization using a redundant path

A third example of a synchronous clocking method proposed by B.K. Ahuja is shown below [figure 6-13] [ref 21]. This structure supplies synchronous clock signals to many points in a system by including a calibration path to mimic the real paths. The clock structure requires a phase lock loop (PLL) and a circuit called a length equalizer. An external pulse train drives the PLL and the output of the PLL drives the global clock generator and the equalizer. Both the clock generator and the equalizer are altered in order to synchronize with the calibration path. Any adjustments by the PLL due to changes in the calibration path are also carried out on the paths to the distant points by way of the equalizer circuit. This design is a more integrated approach and allows all elements of the circuit to remain on the same substrate. It also shows the benefit of using dynamic control not just to measure synchronization but to maintain synchronization. However, this circuit also suffers from the possibility that different paths may drift due the environmental effects or time degradation and also lacks the ability to detect these changes.

6.4) The need for a new clocking method

The impetus for the development of a distributed synchronous clock arose from the nature of the optical interconnection system used to connect the smart pixel arrays. As outlined earlier (see Chapter 3 – Optical Interconnects), the optical system is an uni-directional, closed-ring interconnect. The data within the optical backplane circulates in one direction as it passes through each smart pixel array. The data returns to its origin because the path closes on itself. In this way, any node can communicate with any other node in the system without requiring a bi-directional link. However, because the interconnect closes on itself, there is no “beginning” or “end” to the interconnect path.

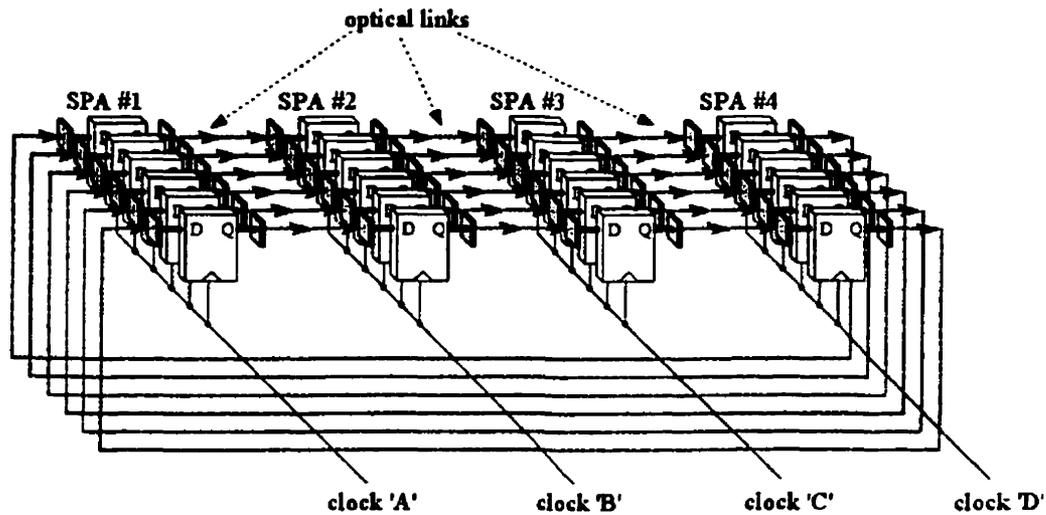


Figure 6-14: A uni-directional ring of DFF

Any point along the path can be considered as the “beginning”. This is why an extremely well synchronized clock must be distributed to all the smart pixel arrays. Each smart pixel array contains several parallel channels of delay flip-flops that must be clocked so that certain logical functions can be performed [figure 6-14].

In order to clarify the timing issues of the uni-directional closed-ring interconnect, a similar architecture can be used to gain some insight. A technique used in most present day high-speed microprocessors is the instruction/data pipeline [figure 6-15] [ref 22]. To use a pipeline, a task is broken into a sequence of several smaller sub-tasks, where each sub-task can be processed very rapidly. By breaking a task into smaller pieces, the overall throughput of the system can be increased. Since each stage in the pipeline can

accomplish its sub-task very quickly, the clock frequency of the computer can be maximized. It is the longest delay between stages of the pipeline that dictates the period of the clock pulses. Using a pipeline, a processor can provide completed tasks at an increased rate. However, the latency between initiating a task and the completion of the task still requires multiple clock cycles.

The similarity between a pipeline used in a processor and the uni-directional closed-loop data path in the smart pixel array interconnect is that both use a sequence of registers with very fast combinational logic between each register. The major difference is that the processor pipeline is built on a single silicon

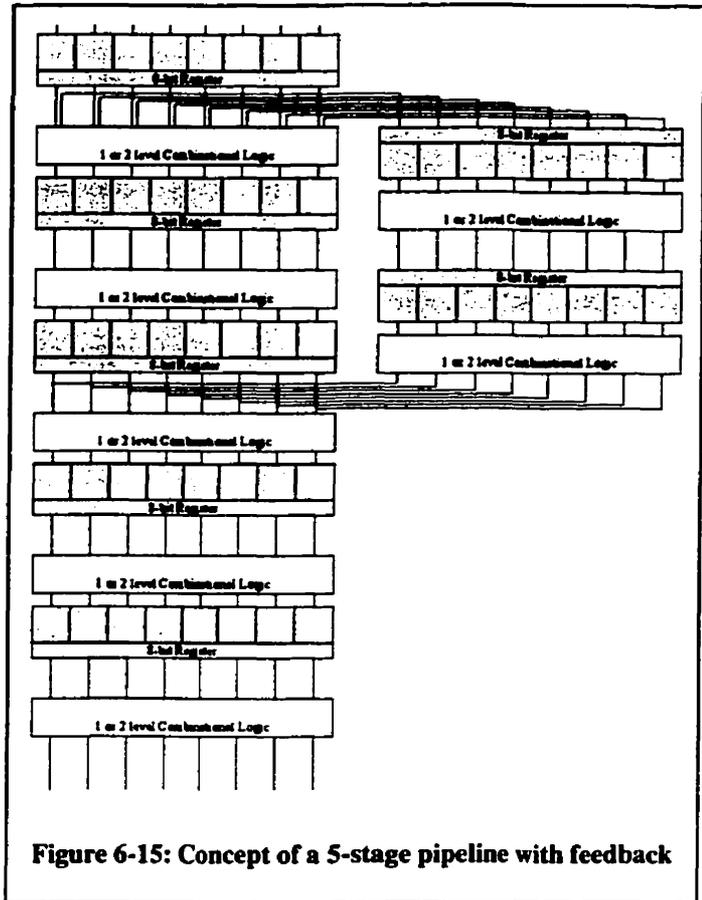
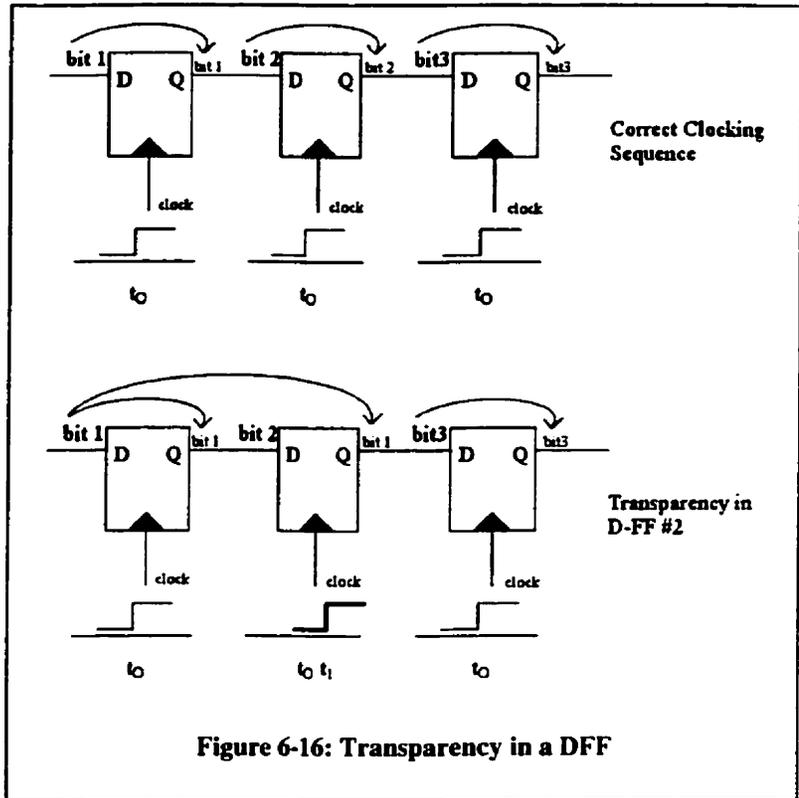


Figure 6-15: Concept of a 5-stage pipeline with feedback

chip whereas the registers in each smart pixel array are on different PCBs and are interconnected using several centimeters of free-space optics. If each register in the pipeline of a processor is not triggered at precisely the same moment, a register can appear “transparent”. A transparent register means that it reacts too late and data from the ‘(i-1)th’ register is passes to the ‘(i+1)th’ register, overwriting the data in the ‘ith’ register [figure 6-16]. The same scenario can occur in the smart pixel array interconnect. Therefore, it is essential that there is very little clock skew among the smart pixel arrays even though relatively large physical distances separate the arrays. A simulation involving four D-flip flops connected in an uni-directional closed-ring shows how a minor clock skew among nodes can result in corrupted data. The D-flip flops (D-FF) were designed using 1.2-micron CMOS technology in the typical master-slave

configuration [figure 6-17]. The D-FF has a set-up time of 0.8-ns, and a clock-to-output time of 1.2-ns.

Each of the four interconnects were simulated using four minimum sized CMOS inverter gates to provide a total delay of 1.1-nsec in each optical path. The circuit also included four 2-input CMOS multiplexers with an average input-to-output delay of 2.3-ns. These multiplexers were used to inject data into the closed-



ring. The multiplexers were then immediately returned to the “straight-through” path. When the four clocks of each D-FF were in perfect phase alignment, the closed-loop unidirectional ring performs properly [figure 6-18]. However, when one of the clocks is

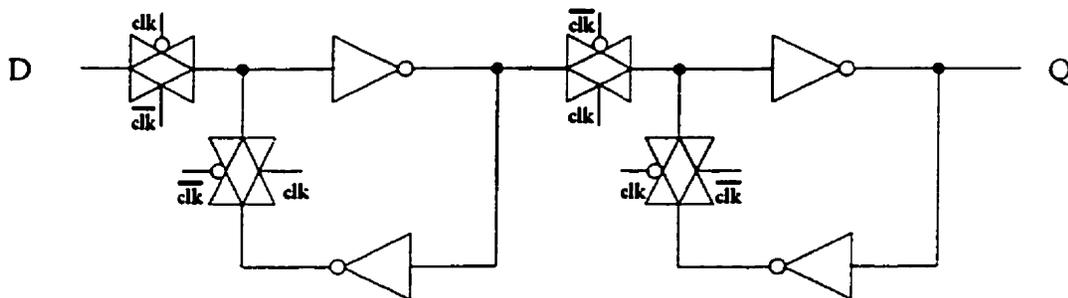


Figure 6-17: A master-slave CMOS DFF

delayed with respect to the others, the loop no longer functions properly and data begins to overlap [figure 6-19]. The maximum skew between clocks for this circuit was approximately ± 150 -psec.

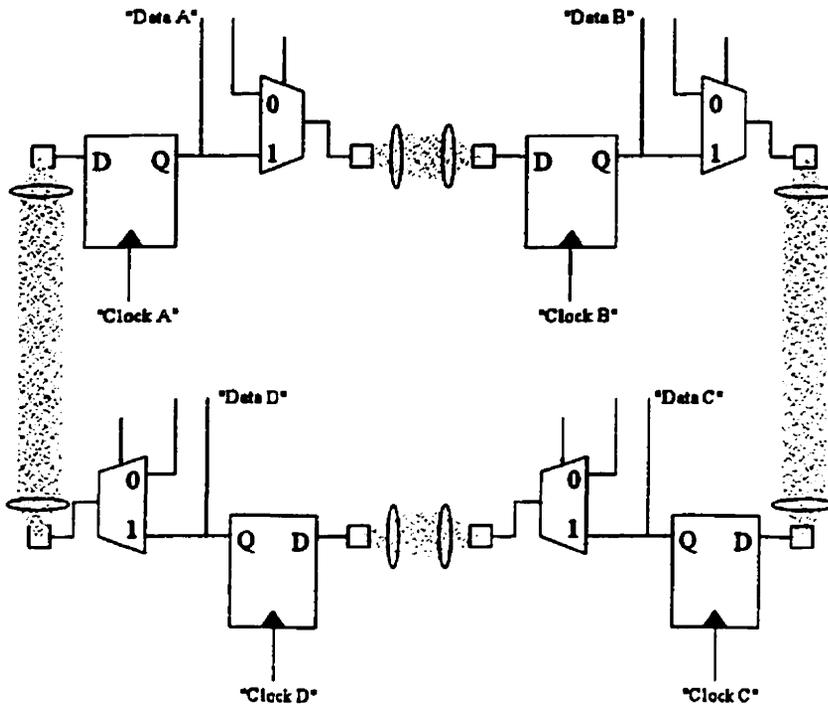
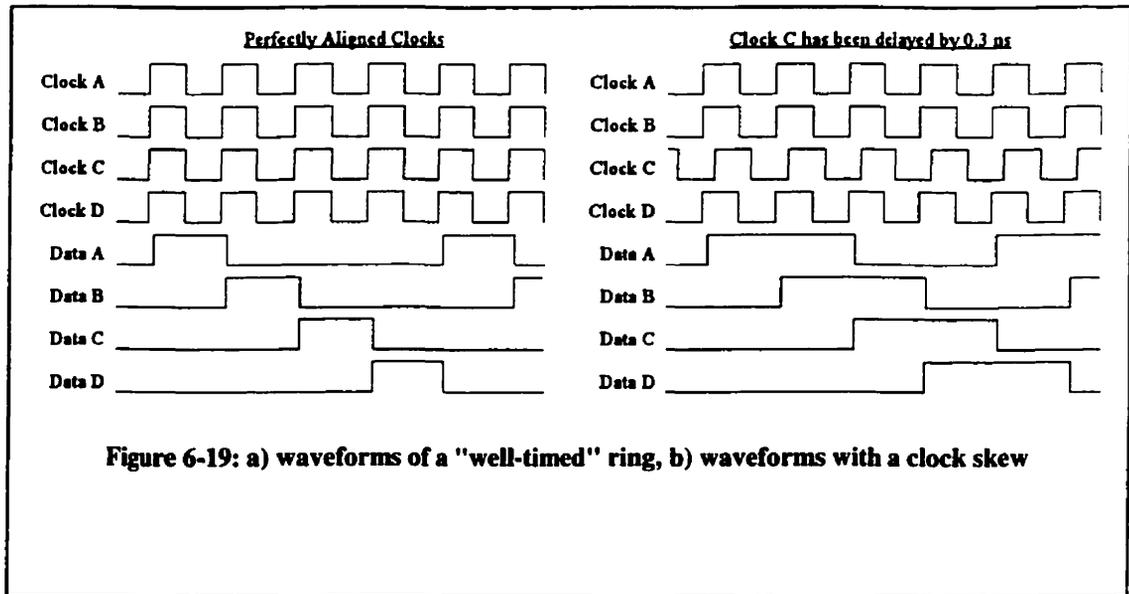


Figure 6-18: Optical/electrical DFF ring



6.5) The development of the distributed synchronous clock

The following subsections describe the chronological development of the distributed synchronous clock, from initial concept to final design. First, the system that this clocking structure targets, as well as a list of desired features for the clocking structure is given. The subsequent subsections provide the development of the clocking system. Each subsection builds on the previous one in order to explore the downfalls and benefits of the particular circuit. Each proposed circuit attempts to fix a problem inherent in the previous circuit, which finally leads to a complete clock synchronization circuit.

6.5.1) The target system for the distributed synchronous clock

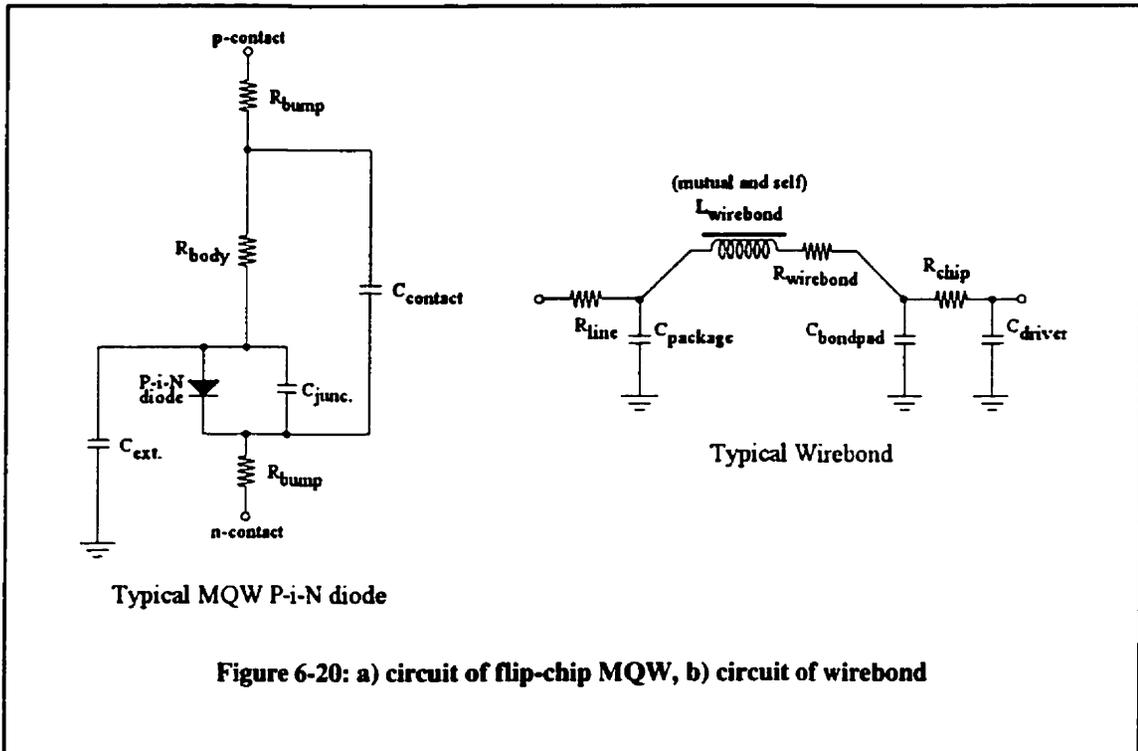
The board-to-board optical backplane was the initial target of the distributed synchronous clock. The optical backplane contains several PCBs, and each PCB has at least one smart pixel array. The size the system may range from a few centimeters up to several meters, and as such, may suffer from large clock skew. The distributed synchronous clock was to have the following properties [Table 6-2]:

- 1) The clock must run at a very high frequency (> 100-MHz).
- 2) It must have virtually zero skew among every node in the system.
- 3) The clock must be scaleable up to any number of PCBs.
- 4) It has to be a fully integrated technology such that all signals remain at the speed of the processor microelectronics.
- 5) It must be fully integrated into the data paths of the interconnect such that separate clock distribution circuitry is not required.
- 6) The clock distribution network must generate its own pulse train without the need for a unique clock generator.
- 7) And no central location for clock-skew control may exist; each node must sense and adjust for variations in clock skew by itself.

Table 6-2: List of properties for the distributed synchronous clock

Some of these features can be immediately satisfied if optical interconnects are used. However, this clocking scheme is not limited to optical interconnects, and there are many ways to satisfy most of these points using electrical interconnects. The free-space optical interconnect inherently satisfies points (1), (4), and (5) by providing a massive number of direct connections via optoelectronic devices directly to the processing circuitry on the chips. Given the large number of point-to-point data connections, several

paths could be reserved for the clock signals without any impact on the connection density of the data paths. The added benefit of the optical interconnect is that it offers a low impedance to the propagation of the optically encoded signals. Circuits on two separate optoelectronic chips can communicate with each other at rates comparable to normal intra-chip speeds. The optical interconnect may also have a very low power consumption compared to high-speed electronic interconnects. The optoelectronic devices offer very small capacitive loads to the microelectronic circuits compared to the typical load of electrical bond pads, wirebonds and trace lines [ref 23] [figure 6-20].



6.5.2) The digital ring oscillator

The digital ring oscillator is a very simple integrated microelectronic oscillator and can be used as the source of a high-speed clock. The ring oscillator is composed of several gain elements connected in positive feedback. A typical example of digital ring oscillator uses single ended CMOS inverters to provide gain for the oscillation. The inverter also provides a 180° phase shift to the input signal and therefore positive feedback is achieved when an odd number of inverters are connected in a ring. This is an unstable configuration

and causes the ring to oscillate [figure 6-21]. The differential pair is another example of an element that can be connected in a ring to produce oscillations. The differential pair can also provide gain for the

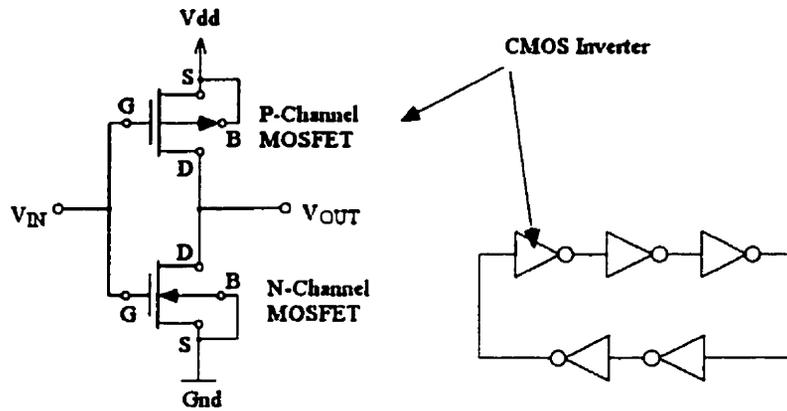


Figure 6-21: CMOS ring oscillator using inverters

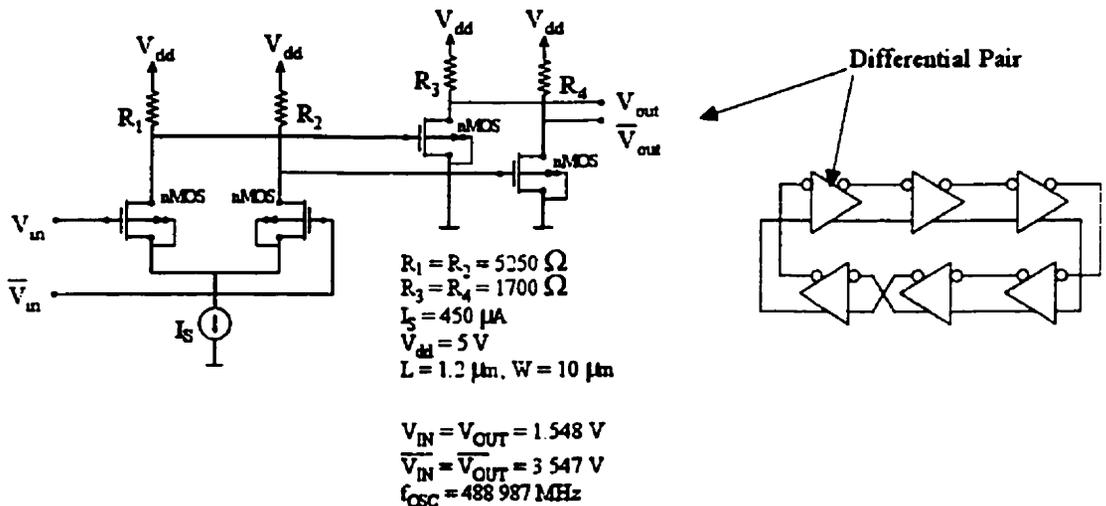


Figure 6-22: Differential Pair Ring Oscillator

oscillation but requires an odd number of cross-coupled interconnects between gain stages to create the positive feedback. This circuit is considerably more difficult to build because a differential pair usually requires a level-shifting mechanism to match the input voltage levels with the output voltage levels. A circuit schematic of a suitable differential pair circuit, including the level-shifting stage, is shown above [figure 6-22].

The perturbation in the digital ring oscillator is manifested as a pulse or “event” which travels around the ring reversing the state of each node as it propagates [figure 6-23]. A reasonably good analogy of this travelling “event” is the mechanical wheel and spoke [figure 6-24]. As the wheel turns, the spoke passes each element (in this case the four boxes) one after the other. The spoke is equivalent to the “event”, and the frequency of oscillation is obtained from two revolutions of the spoke. Two revolutions are required because the event changes the state of the node (i.e.: one period consists of a low to a high and back to a low transition).

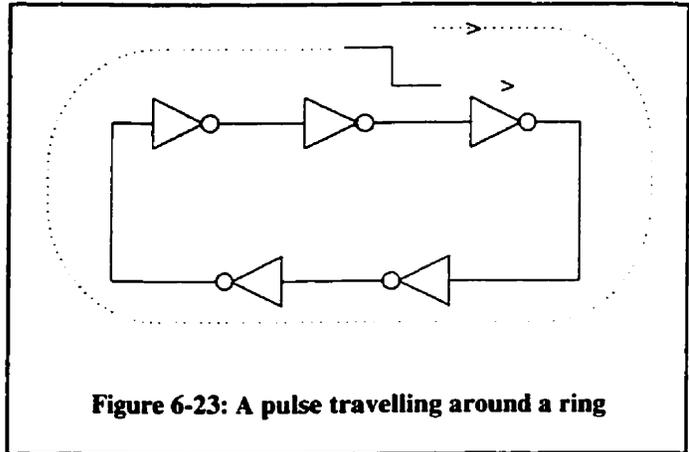


Figure 6-23: A pulse travelling around a ring

The number of elements in the ring, the speed of each element, and the “time of flight” between elements are all aspects that can affect the frequency.

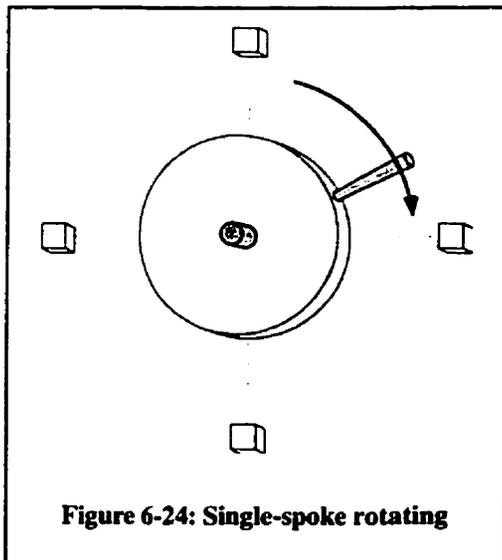


Figure 6-24: Single-spoke rotating

The number of elements in the ring, the speed of each element, and the “time of flight” between elements are all aspects that can affect the frequency.

A ring oscillator with a variable frequency output is called a voltage-controlled oscillator (VCO) and is one of the main components in a phase-locked loop (PLL). A typical VCO is shown below [figure 6-25]. The bias to the inverters can be altered in order to change their speed of operation. A VCO

usually has only one tap-point and the frequency generated at this point is generally very stable. Varying the “quasi-dc” control line can alter the frequency and thus a voltage-to-

frequency transfer function is obtainable. The VCO converts a 'dc' control voltage into a frequency. The VCO is ideal as a frequency reference, but not as a reference of absolute phase. Since frequency is the derivative of phase and its integration produces an unknown constant, no information about phase can be obtained from the frequency alone. The constant "K_{VCO}" is simply the slope of the line from the graph of voltage bias versus angular frequency (see APPENDIX A).

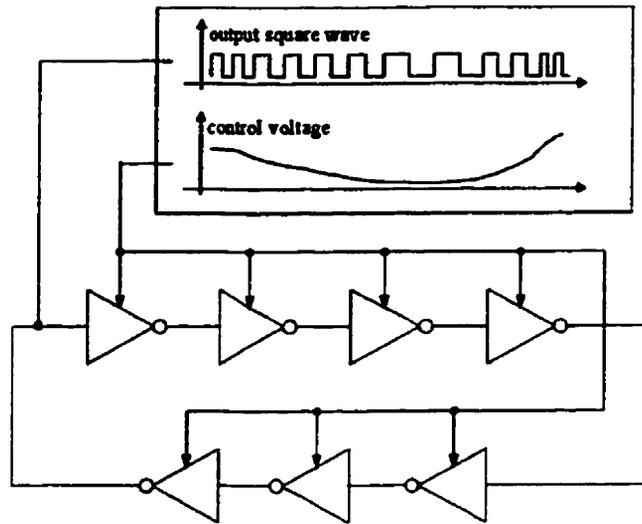


Figure 6-25: A typical VCO

Eqn. 6-1 a,b,c):

$$\omega(v_{dc}) = \text{function of } (v_{dc})$$

$$\omega(v_{dc}) = K_{VCO} \cdot v_{dc}$$

$$\omega(v_{dc}) = \frac{\partial \theta(v_{dc})}{\partial t} = K_{VCO} \cdot v_{dc}$$

The ring oscillator was the basis for most of the initial work on the distributed synchronous clock and elements of this circuit are present throughout most of the following subsections. This circuit combined with the properties of the optical interconnect create the second step in the evolution of the distributed synchronous clock; this is called the optical ring oscillator.

6.5.3) The optical ring oscillator

The combination of the digital ring oscillator and the uni-directional optical interconnect allow for a unique structure called an optical ring oscillator (ORO). The ORO is composed of banks of active electronic elements, such as CMOS inverters, and passive optical interconnects between these banks. The ORO is identical to the digital ring oscillator except that the interconnect lines between some electronic elements are optical beams of light. As indicated in a previous chapter (see Chapter 2 – Architecture), the *transparent-state* of a smart pixel could permit the oscillator to function at very high rates even though the signals travel between PCBs. The ORO possesses at least one the characteristics necessary for the generation and distribution of high-speed clocks distributed over large distances. By implementing the correct number of electronic inversion within the smart pixels, the uni-directional closed-ring optical interconnect can implement an optical ring oscillator. A simplified picture of the ORO, describing four nodes (or smart pixels) connected in a closed optical ring, is given [figure 6-26]. One of the four nodes has an even number of CMOS elements due to the controlling multiplexer, which makes an odd number of inverters for the entire ring. Note that the ORO can be used as a voltage controlled optical ring oscillator (VC-ORO) by adjusting the bias of the elements at any one node thereby affecting the overall frequency of the ring.

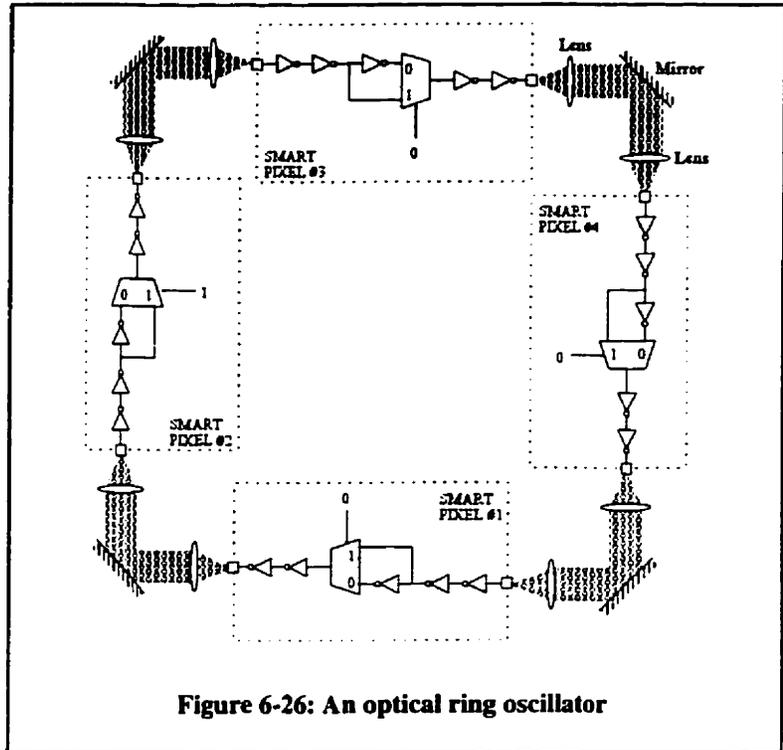


Figure 6-26: An optical ring oscillator

Architecture), the *transparent-state* of a smart pixel could permit the oscillator to function at very high rates even though the signals travel between PCBs. The ORO possesses at least one the characteristics necessary for the generation and distribution of high-speed clocks distributed over large distances. By implementing the correct number of electronic inversion within the smart pixels, the uni-directional closed-ring optical interconnect can implement an optical ring oscillator. A simplified picture of the ORO, describing four nodes (or smart pixels) connected in a closed optical ring, is given [figure 6-26]. One of the four nodes has an even number of CMOS elements due to the controlling multiplexer, which makes an odd number of inverters for the entire ring. Note that the ORO can be used as a voltage controlled optical ring oscillator (VC-ORO) by adjusting the bias of the elements at any one node thereby affecting the overall frequency of the ring.

With the ORO, it is possible to generate a fixed frequency common to every node in the backplane. It is also possible for the ORO to serve as both a clock generation and

distribution mechanism for the optical interconnect. However, because there is no reference phase common to all nodes, the ORO cannot be used as a synchronous clock source.

6.5.4) The multiple tap-point ORO

If the ORO is slightly altered to provide a 'tap-point' at each node in the ring, then each node can receive the same frequency square wave. Each tap-point produces an identical frequency but out-of-phase from the previous tap-point in the ring by the delay

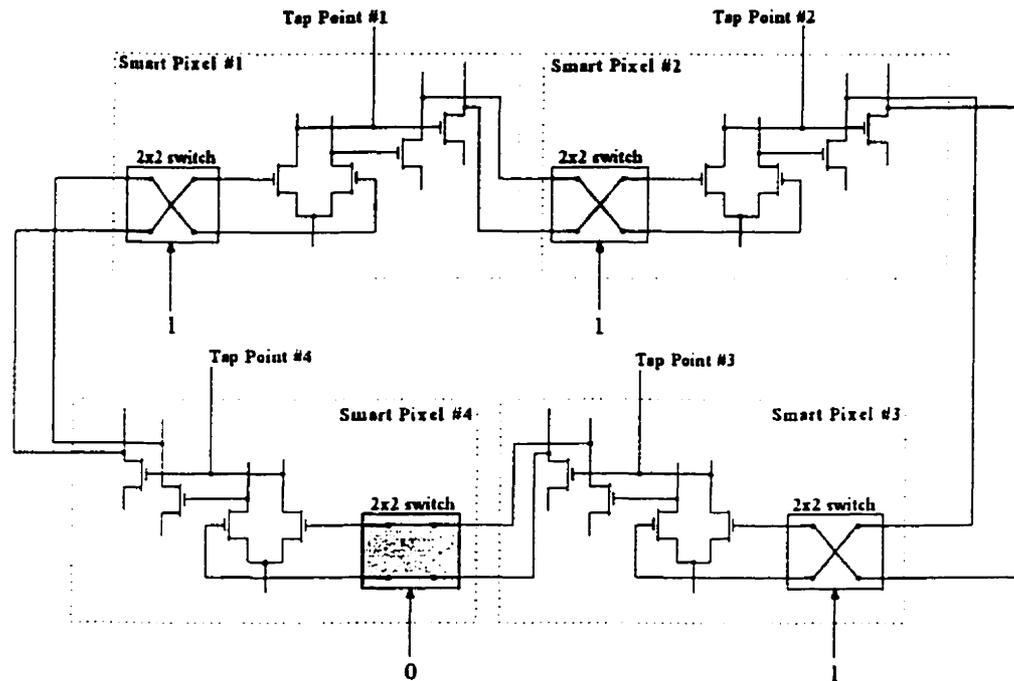


Figure 6-27: A 4-node differential pair ORO

between the nodes. Allowing each node in the ORO to be constructed from differential pairs, an even number of nodes can be used where only one path must be cross connected in order to form the positive feedback. An example of a four-node differential ORO is given [figure 6-27]. If it is assumed that the delay between each of the four tap-points on the ring is equal, a PLL can be used to lock onto the ORO frequency and produce a frequency eight times higher than the fundamental frequency of the ORO [figure 6-28]. The PLL structure simply requires a divide-by eight register in the feedback path to boost the frequency (see APPENDIX A). A set of typical waveforms for the multiple tap-point

ORO are given [figure 6-29], and show the ideal case when all outputs of the nodes are perfectly aligned in phase.

Unfortunately, proper synchronization of all nodes assumes that the delay between tap-points is identical. However, it is likely that some mismatches in delay will occur between tap-points and therefore there will be an added phase error. The phase error between nodes cannot be adjusted because the frequency of the ORO is the only observable variable at each node. Since phase is the time integration of frequency, absolute phase information is lost when the

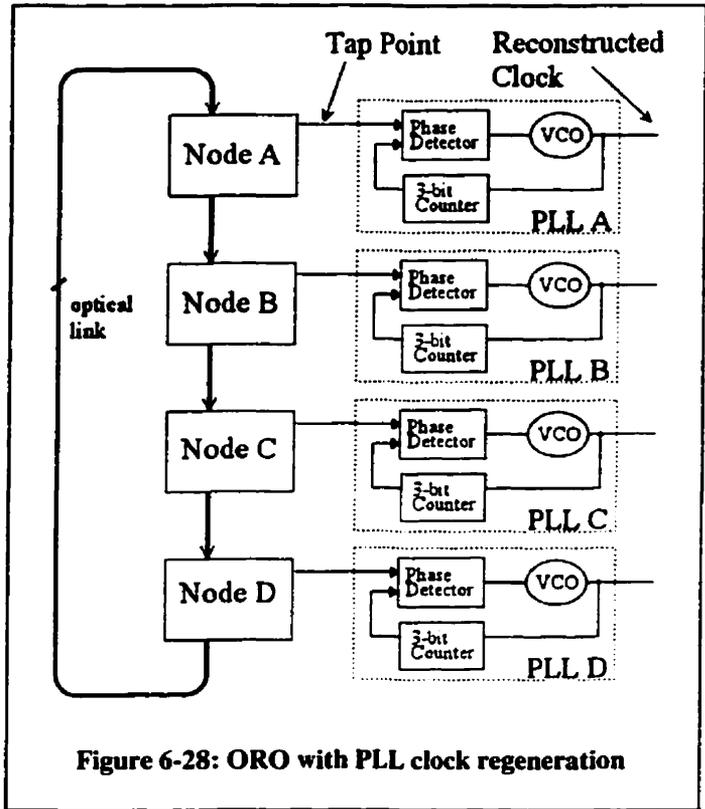


Figure 6-28: ORO with PLL clock regeneration

initial conditions are not known. Described in another way, the total delay around the ring determines the overall period of the ORO. However, some parts of the ring may be faster

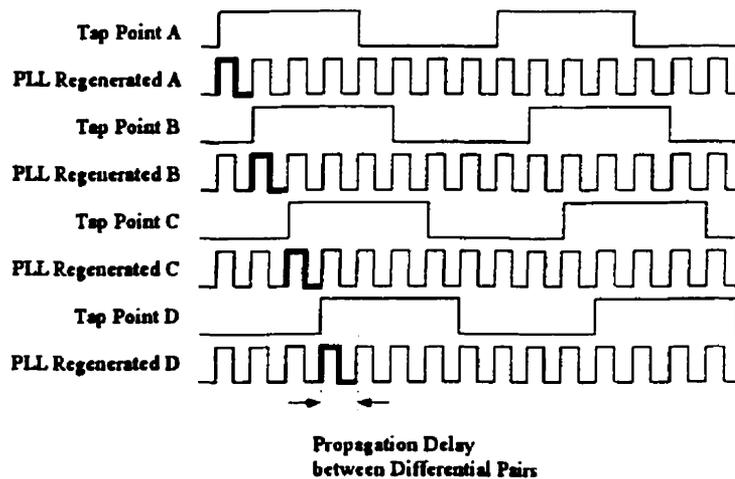


Figure 6-29: Waveforms for ideal ORO

than others due to effects such as thermal gradients, process variations, or differing biasing conditions. The differences in propagation delay would shift the edges of the regenerated waveforms of the PLLs at each node such that they no longer coincided in time. Each node would have radically different duty cycles

and different relative phase delays [figure 6-30].

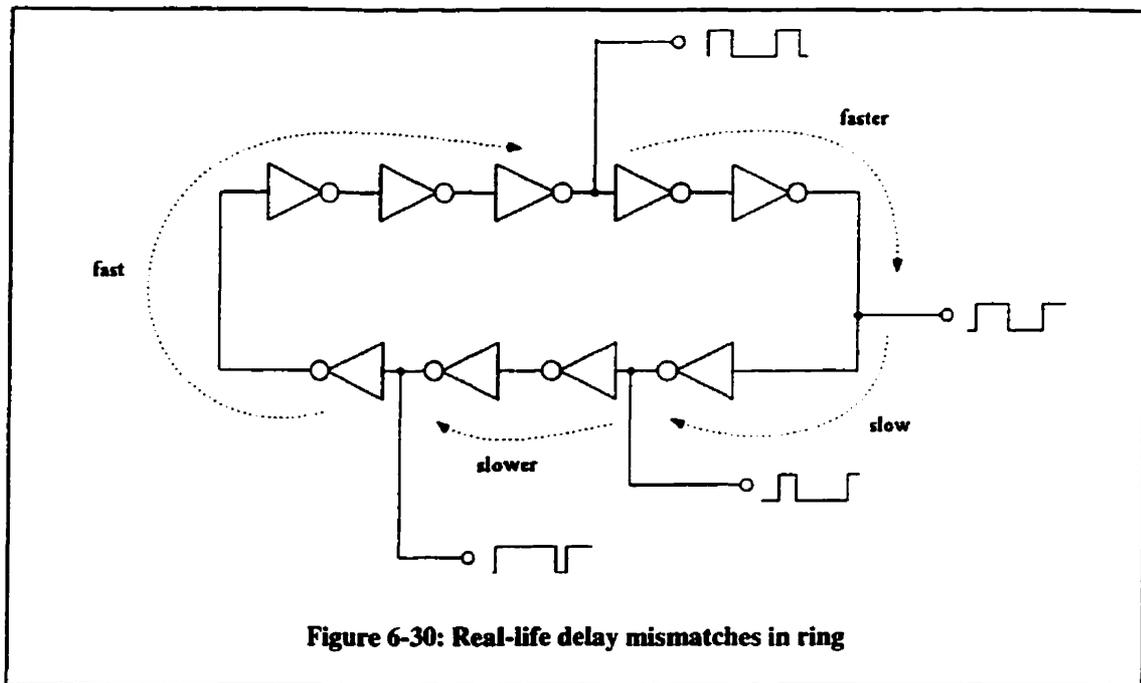


Figure 6-30: Real-life delay mismatches in ring

The only way that the multiple tap-point method would work is if the propagation delays between each pair of nodes were equal. In the next subsection, the use of a global control mechanism will be discussed in order to maintain a fixed delay between nodes. The global balancing mechanism would function in a way analogous to a bus controller on an electronic backplane by assuming control of everything. The clock controller would be responsible for calculating and adjusting the delays between pair of nodes so that the clocks would remain in phase at each tap-point.

6.5.5) A global clock control mechanism

One of the more straightforward methods to force all nodes to maintain synchronous clocks would be to use some form of global control. Information about the delay between each pair of nodes in the ORO could be transmitted to a central controller, and the controller could issue instructions to re-balance the clock signals. Alternatively, information about the delays in the ORO could be disseminated to all other nodes in the ring, and each node could re-balance itself [figure 6-31]. However, there are several problems with distributing this type of information, the first problem is more of a

philosophical question concerning the global control of a central controller. If a central controller must be built, all the connections to all points of the systems must be made. If this was the case, then the same resources could be used to implement a more traditional clocking scheme such as a binary-tree with much less difficulty. The second issue concerns the dissemination of all the information to all the nodes. In order for all the nodes to access all the information about the network, the network would become exponentially more congested, of order $O(n^2)$, with signals which carry only information about delays. An issue common to both global control schemes mentioned above is that the transmission lines carrying the information are themselves unknown delays. If the delays between nodes vary with time (i.e.: due to thermal cycling) then a continual lag between the actual state of the system and the controlling signals could occur.

Based on the drawbacks outlined here, it was decided that the global control mechanism was not a suitable method to control path delay variations. The method of local control at each node seemed to be the best choice. Each node must be able to sense phase error, and adjust for it, in order to achieve global synchronization. There is no need for a master control or a method of distributing information about delay though the system. A method that can provide a way for all nodes to simultaneously sense phase error is given in the next subsection.

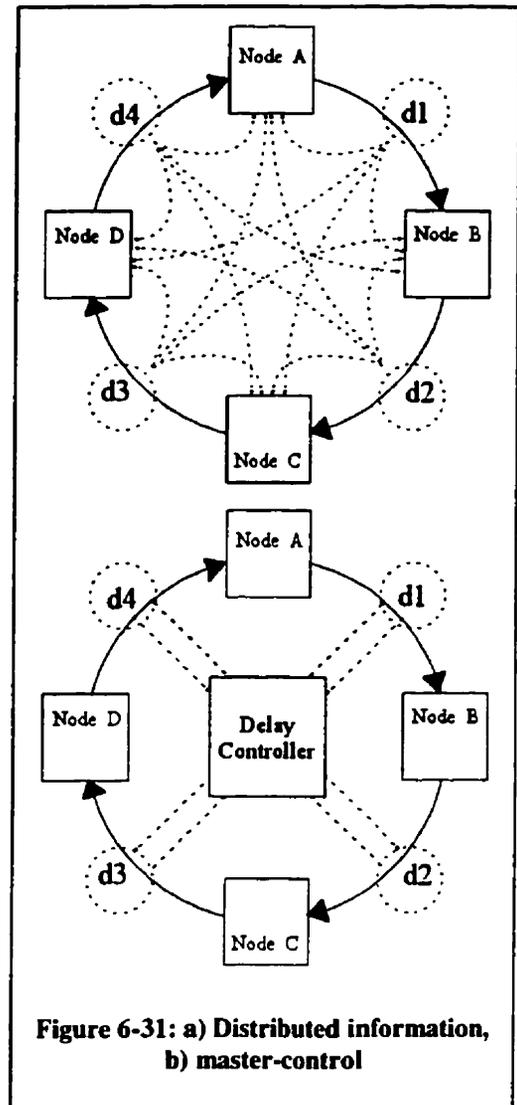


Figure 6-31: a) Distributed information, b) master-control

6.5.6) Spatially separated multiple phase generation

If the ORO is used in its present state to distribute a clock signal to all nodes, a single pulse would propagate from node to node in succession. In order to use the

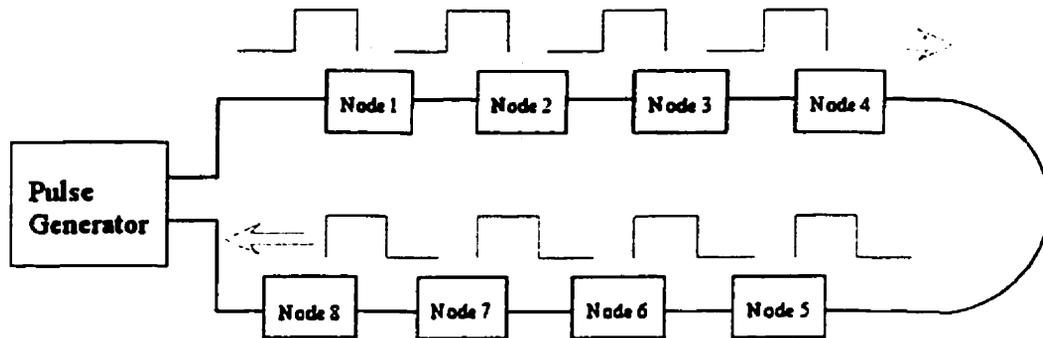


Figure 6-32: Concept of a multiple-pulse generator

concepts of basic control theory [ref 24] each node must compare two similar quantities to produce an error signal that in turn will adjust some parameter to force the error to zero. This suggests that the ORO circuitry must be modified to produce multiple pulses so that each node receives at least two pulses in order to measure a difference. The concept of a spatially separated multiple pulse generator, capable of producing the same number of pulses as there are nodes, is shown here [figure 6-32].

A spatially separated multiple pulse generator must produce pulses, separated in time and space, such that every node receives a pulse simultaneously. The mechanical wheel and spoke analogy is again useful to help visualize the circuit [figure 6-33]. As the wheel turns, all the spokes pass a box simultaneously (assuming that the distances between

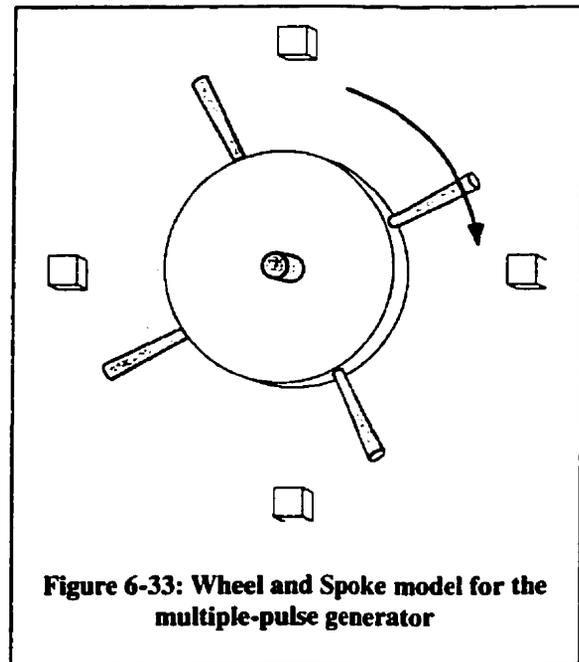


Figure 6-33: Wheel and Spoke model for the multiple-pulse generator

spokes are equal and fixed). No central control is required because the spokes (or pulses) are inherently spatially and temporally separated and are able to maintain the correct

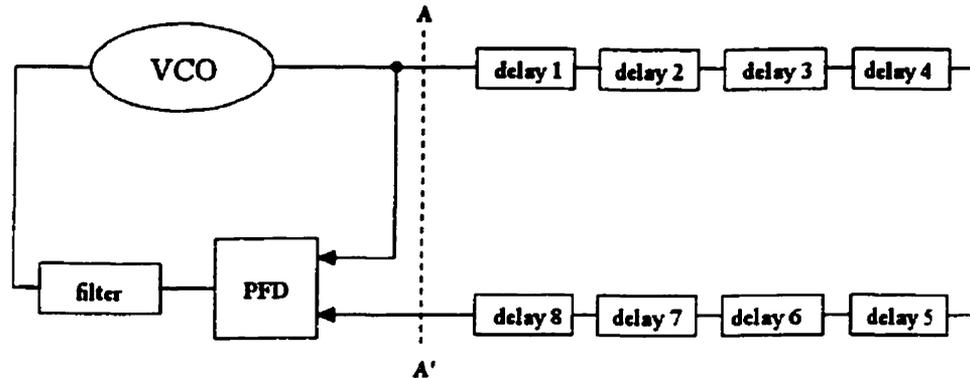


Figure 6-34: First iteration of the multiple-pulse generator

number of pulses for a given number of nodes.

The following circuit was one of the first attempts at realizing this concept [figure 6-34]. This circuit attempts to maintain exactly eight pulses within the eight delay units using a PLL (the VCO, filter, and PFD). The error signal is generated using the output of the VCO and a delayed version of the same output. The PFD is a digital circuit that allows this comparison. If the delay of each unit in the delay line is 10-nsec and the center frequency of the VCO is 100-MHz, the total delay around the delay line is 80-nsec and the VCO produces eight cycles within the delay line. Unfortunately, this circuit is not stable and cannot guarantee that eight pulses exist in the eight nodes. If the oscillation of the VCO was assumed to be 87.5-MHz, exactly seven cycles would exist in the delay line separated by 11.4-ns. If the oscillation of the VCO was 112.5-MHz, there would be exactly nine cycles in the delay line separated by 8.89-ns. Each of these frequencies is reasonably close to the center frequency that it is possible for a single VCO to produce any of them. Therefore, this circuit would not be suitable to produce a predictable number of pulses within the delay line.

A slightly alteration to this circuit allows the number of pulse to be correlated to the number of nodes. This circuit [figure 6-35] provides two delay lines that are identical. One of the delay lines is directly attached to the VCO and is called the “fast” delay line, where the letter ‘F’ is used to represent “fast”. The other delay line is attached to the most

significant bit of a 3-bit synchronous counter and is called the “slow” delay line, where the letter ‘S’ is used to represent “slow”. The output of the VCO drives the clock of the 3-

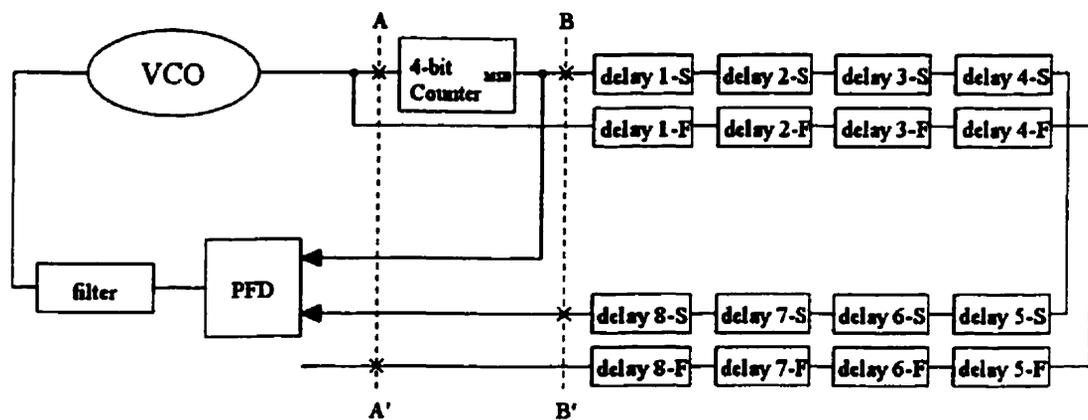


Figure 6-35: Second iteration of the multiple-pulse generator

bit counter and produces a frequency eight times slower. The 3-bit counter produces the same number of pulses as nodes in the “fast” delay line.

The PFD produces an error signal based on the output of the 3-bit counter and the delayed version of the same signal along the “slow” delay line. If it is assumed that the circuit is composed of eight nodes each with 10-nsec of delay and that the VCO is operating at 100-MHz, the “fast” delay line will carry eight pulses separated by 10-nsec and the “slow” delay line will carry one pulse of 80-nsec. Since the PFD circuit is connected to the beginning and the end of the “slow” delay line, the PFD measures a rising edge pulse at both the beginning and the end of the “slow” delay line due to the 3-bit counter. While producing exactly eight pulses in the “fast” delay line during the same interval. If the VCO was initially at 87.5-MHz and seven pulses exist in the “fast” delay line, the frequency of the “slow” delay line would be 10.93-MHz (after a divide by eight in frequency) hence, only 0.875 of a pulse would exist in the “slow” delay line and this does not satisfy the PFD. The next valid operating frequency for the VCO is 200-MHz where sixteen pulses separated by 5-nsec would exist in the “fast” delay line and two pulses separated by 40-nsec would exist in the “slow” delay line.

The circuit in [figure 6-34] can allow the VCO to oscillate at 100-MHz, 87.5-MHz, or 112.5-MHz and still satisfy the requirements of the PFD circuit. The circuit in [figure 6-35] requires a frequency change from 100-MHz to 200-MHz in order to satisfy

the first and second harmonics of the delay lines. Even though a 200-MHz frequency allows the circuit to function, a VCO with a center frequency of 100-MHz is unlikely to settle on a 200-MHz solution primarily because of its limited operating range.

The number of nodes is directly related to the number of bits in the counter, for 8 nodes, a 3-bit counter is required. A 4-bit counter implies 16 nodes and a 5-bit counter implies 32 nodes. If the number of nodes required is not a multiple of 2, then the counter must be made more complicated. If 10 nodes are required, a 4-bit finite state machine must be implemented where the reset signal is produced every 10 cycles. If an odd number of nodes are required, then a similar finite state machine must be used, but the PFD must be insensitive to NON-50/50 duty cycle clock periods.

The fact that two delay lines must be used in order to propagate two signals around the system in the circuit as shown in [figure 6-35] might be thought of as a waste of resources and bandwidth. However, there are methods which can allow the same transmission medium to simultaneously carry both counter-propagating signals between nodes [ref 25]. Also, data encoding techniques can be used to transmit both the “fast” and the “slow” signals along the same medium; hence the number of delay lines is reduced to one. However, for clarity in the description of the circuit, each transmission path will carry only one signal. There is an obvious benefit when using a single transmission medium because any small variation in delay that could occur between two separate lines is no longer possible. If a purely digital encoding technique is not possible, a analog method of sending carrier frequencies could also be employed. Two modulated carriers could be transmitted and then separated using bandpass filters at each node in the system. The demodulated signals could then act in the same manner as the digital pulses described above. This method would obviously involve a significant amount of analog electronics, but is still a reasonable way to use a single medium.

The specific details of PLL lock-in time and stability are very complicated and involve many non-linear processes. These topics are not dealt with in this thesis; however, most texts on PLLs do investigate these issues [ref 26]. Without directly addressing the issue of how power supply jitter alters VCO frequency or how filter noise

adds an unwanted dynamic to the control system, there are several simple techniques that can help ensure proper operation of the circuit proposed above.

One of the main difficulties associated with the circuit above [figure 6-35] is the appreciable delay around the entire delay line. This significantly long delay may cause the PLL to force the VCO out of its operating region upon start-up. If the PFD reference input receives a signal, but the second PFD input does not, the VCO may be forced to shut off due to the increasing error signal. To correct this problem, a mechanism which allows the system to initialize the delay line can be included such that the VCO would remain at a fixed frequency for a specific amount of time until the delay line has been filled with pulses. Lock-in problems may also be avoided if a couple of additions to the circuit are made. To aid in the stability of the circuit the filter properties of the PLL must be adjusted such that the filter bandwidth is very narrow and thus the circuit has a long settling time. This will help integrate over a longer interval of input signal before changing the error signal.

6.5.7) The counter-propagating multiple pulse generator

Thus far, the circuit developed in the previous section produces the same number of pulses as there are nodes in the system. However, as mentioned earlier, if a node is to adjust some parameter within itself, it must be provided with at least 2 independent signals with which to measure a difference. This is the basic mechanism behind all control systems such as the PLL. By modifying the spatially separated multiple pulse generator circuit above [figure 6-35] to include both a clockwise and a counter-clockwise delay line, a counter-propagating spatially separated multiple pulse generator can be constructed. This circuit [figure 6-36] shows how two multiple pulse generators can be combined such that only one reference oscillator is used.

The mechanical wheel and spoke analogy can be used again to describe the concept of two counter-propagating multiple pulse generators. It is assumed that there are 4 nodes in the system. The wheel labeled 'A' has four spokes each separated by 90° and is rotating clockwise. The wheel labeled 'B' also has four spokes each separated by 90° and is rotating counter-clockwise. When each pair of spokes aligns at each of the four nodes, there is zero error and a synchronous event has taken place.

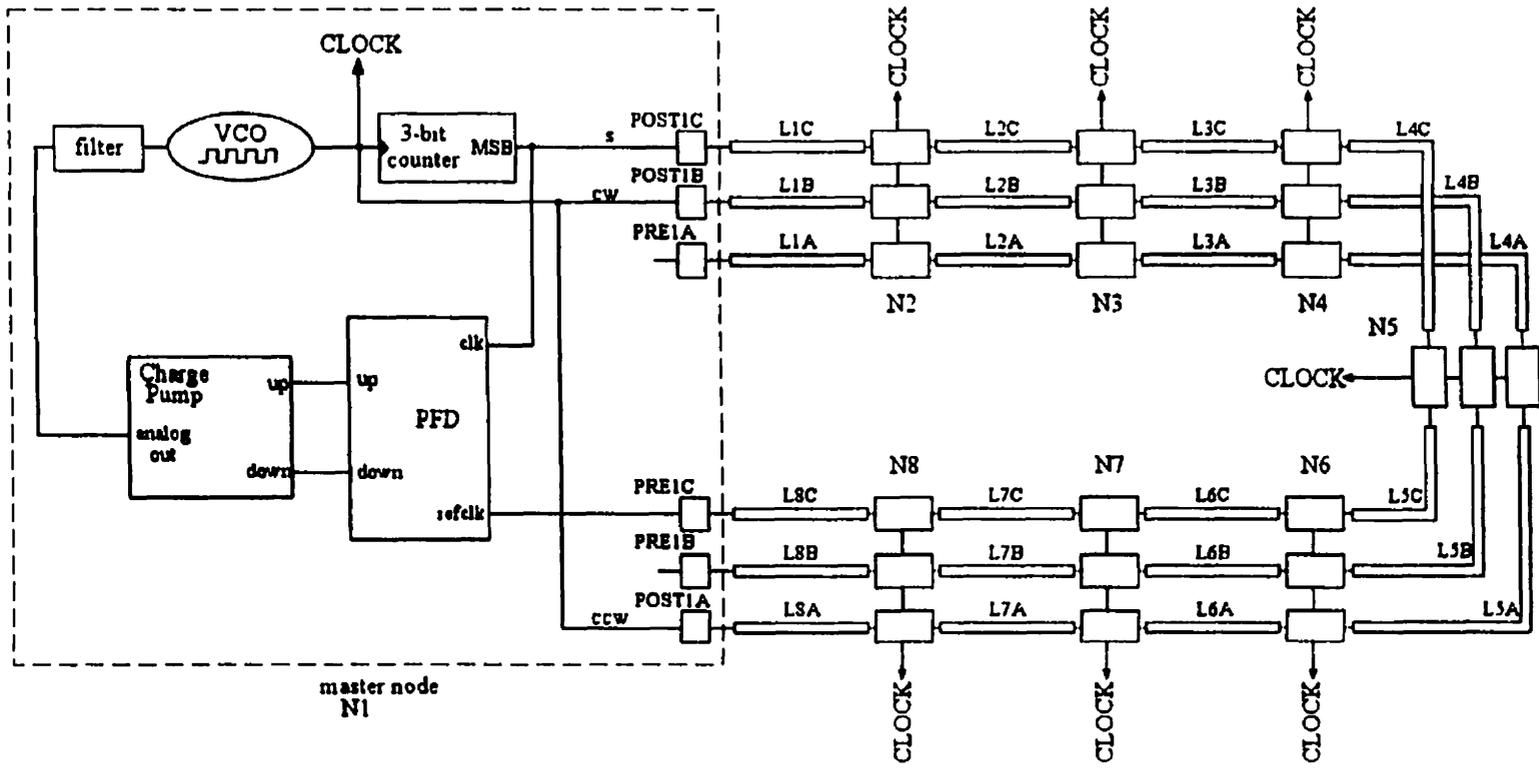
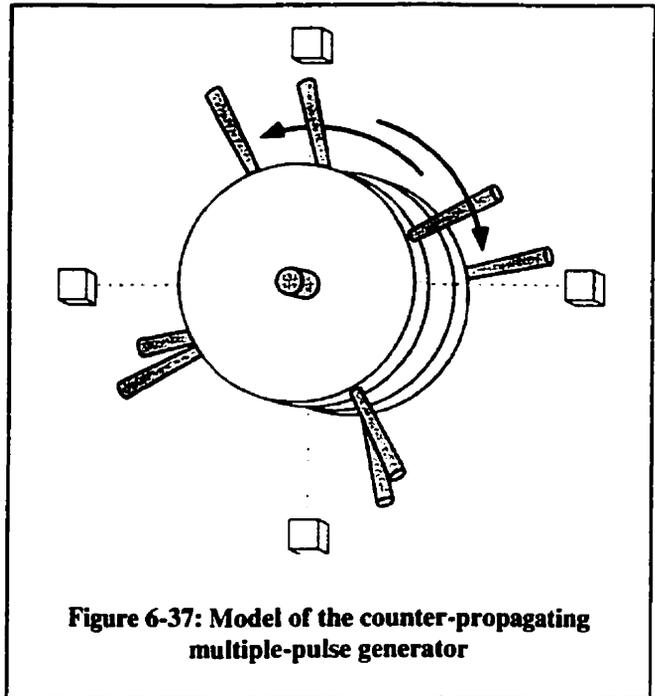


Figure 6-36: The counter-propagating, multiple-pulse generator

However, there are several conditions that must be met for the counter-propagating multiple-pulse generator to work properly. First, all delays between adjacent nodes in both rings must be equal and constant. For example, if one of the 4 spokes on the wheels [figure 6-37] was at an angle other than 90°, it would be impossible for a simultaneous arrival of pairs of spokes at each node. This is analogous to a circuit in terms of the propagation delay between nodes. The propagation delay must be equal for each pair of adjacent nodes. The second condition for proper operation is that both wheels must rotate at the same speed. If this is not the case, the absolute position of spoke alignment will change over time; this is similar to the beat-frequency of two sinusoidal waves with slightly different frequencies. The final requirement is that the pulses in both rings must propagate in opposite directions with respect to each other. This is perhaps the most



subtle and most important aspect of the design. If both signals follow the same path, it is possible to satisfy the first two conditions. However, if a series of pulses with a fixed repetition rate (such as a square-wave) is used in two “same-direction” paths, there is no

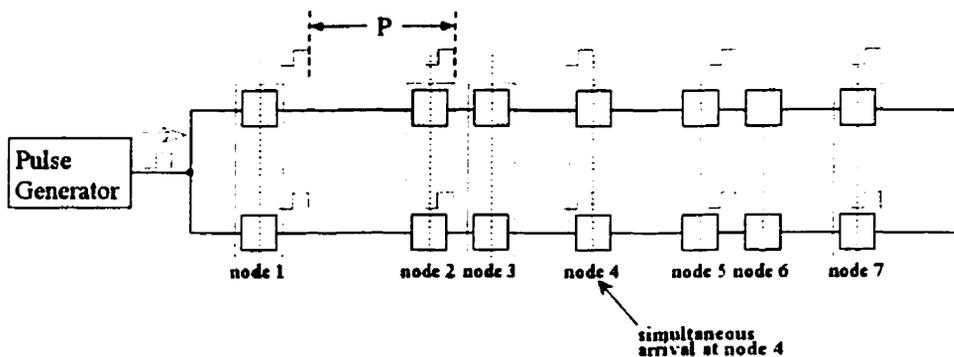


Figure 6-38: Multiple pulses with regular period travelling in the same direction

subtle and most important aspect of the design. If both signals follow the same path, it is possible to satisfy the first two conditions. However, if a series of pulses with a fixed repetition rate (such as a square-wave) is used in two “same-direction” paths, there is no

guaranteed that all the pairs of pulses will arrive simultaneously at each node [figure 6-38]. When the paths of the pulses are opposite to each other, the only way that all the pairs of pulses will simultaneously meet within each node is to guarantee that the delay between any two adjacent nodes is equal [figure 6-39]. This argument is presented as the ultimate objective of these circuits and a limited proof is presented in section (6.6.1).

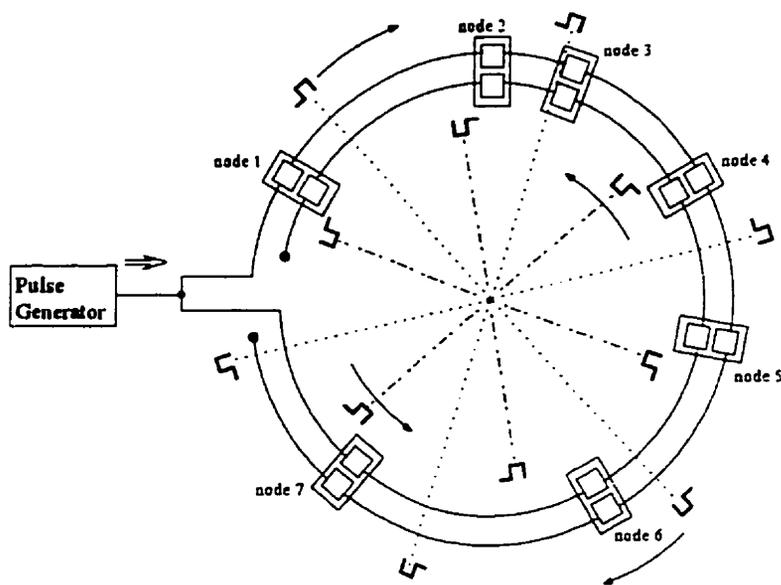


Figure 6-39: Multiple pulses with regular period travelling in opposite directions

6.5.8) Distributed local control

The distributed synchronous clock was constructed because minor skew of clock pulses among multiple points produced register transparency problems for many synchronous systems. This was especially true when the system was physically large such as an optical backplane with many PCBs. To correct for this skew, a method of local control was decided upon. Each node in the system would help balance the entire system without requiring any global knowledge of the state of the system. Each node would be partially responsible for the synchronization of the entire network by slightly adjusting its own propagation delay. An analogy of local control involves the flow of cars on a highway. A method to eliminate traffic-jams on a highway requires that every car be equipped with two sensors, one in the front and one in the back, which it uses to detect

the distance of the cars immediately ahead and behind. Given a predetermined set distance between cars, each car can adjust its speed in order to track to this reference distance. If all the cars on the highway behave in the same way, an orderly and fast flow can be maintained. A paper by B. Bamieh et al. entitled "Distributed Control of Spatially-Invariant Systems" seems to indicate that this solution is feasible given certain boundary conditions [ref 27].

A similar method of local control can be applied at each node of the counter-propagating multiple pulse generator described above [figure 6-36]. Since each node has been equipped with two paths propagating pulses in opposite directions, a method of measuring the arrival of pulses with respect to a certain reference point in each node is possible. The mechanical analogy can once again be used to describe the local control mechanism. The wheel and spoke model [figure 6-40] shows a detailed view of a single node with two spokes (or pulses) arriving from opposite directions. One spoke moves in the clockwise direction, while the other spoke moves in the counter-clockwise direction. As the spokes rotate, they must pass each other close to some "reference line".

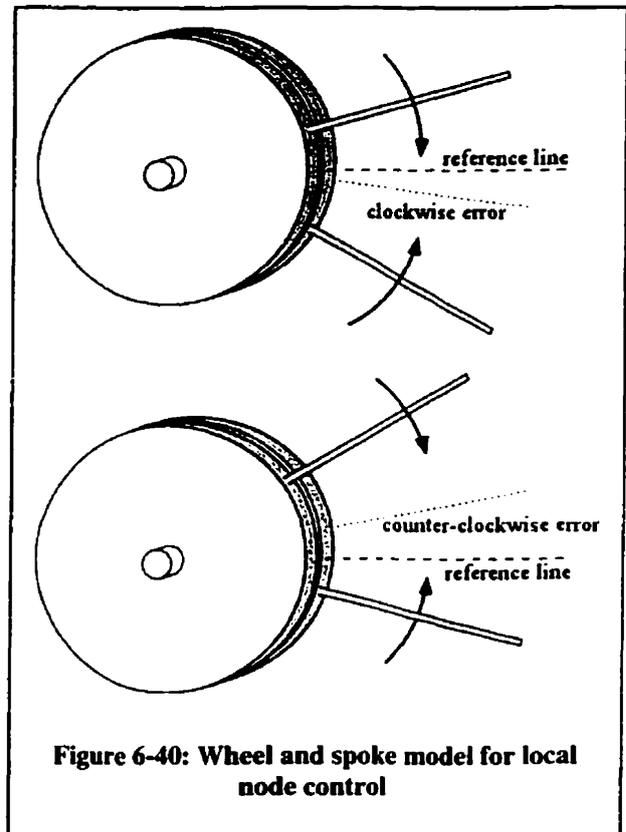


Figure 6-40: Wheel and spoke model for local node control

There are three possibilities that can occur. The spokes can meet at the reference line, before the reference line (counter-clockwise error), or after the reference line (clockwise error). These errors can then be corrected if some form of control is used to actuate minor propagation delay adjustments. The following circuit [figure 6-41] describes one implementation of the local control mechanism described by the wheel and spoke analogy. In the figure, two pulses are shown propagating from opposite directions and heading towards the reference line. The delay line labeled "fast" clockwise has pulses

travelling from left to right. The delay line labeled “fast” counter-clockwise has pulses travelling from right to left.

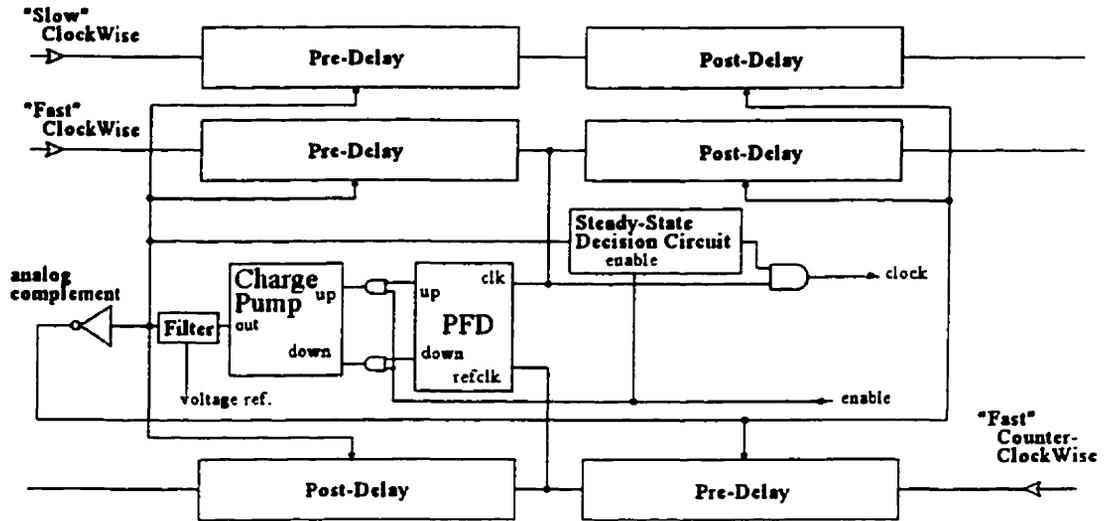


Figure 6-41: A local control node circuit

The points within the local control node where the PFD is connected are equivalent to the reference line. The PFD, the charge pump and the filter are used to produce an error signal that can adjust the internal delays of these paths. If both pulses arrive at the same time at the detection points, then no action is taken. When the clockwise pulse arrives earlier or later than the counter-clockwise pulse, the delay lines preceding the detection points of both delay lines are altered. The following table outlines how the adjustments are made with reference to [figure 6-41] [Table 6-3].

	<i>Pre-Delay</i> <i>“Fast” CW</i>	<i>Pre-Delay</i> <i>“Fast” CCW</i>	<i>Post-Delay</i> <i>“Fast” CW</i>	<i>Post-Delay</i> <i>“Fast” CCW</i>
CW equals CCW	no action	no action	no action	no action
CW leads CCW	delay increased	delay decreased	delay decreased	delay increased
CW lags CCW	delay decreased	delay increased	delay increased	delay decreased

Table 6-3: Conditions on local node circuit variable delays

An extremely important feature of this circuit is that it is a *balancing circuit*. The total delay through both the “fast” clockwise path and the “fast” counter-clockwise path

must remain constant within a node. If a pulse is sped up on the way in, it must be slowed down on the way out. If the absolute delay of a path were changed, then the local control would be either increase or decrease the total delay of the system. What tends to happen is that one path is continually sped-up and the other is continuously slowed-down until the system simply shuts off.

The third path in [figure 6-41], called the “slow” clockwise path, is part of the “slow” delay line first introduced in [figure 6-35]. It is used to carry the low frequency pulse train and is responsible for maintaining the correct number of pulses in the system. As will be shown in the next section, the total delay through a node must not change, however, the delay between the detection point pairs within adjacent nodes is allowed to change in order to satisfy the overall delay balancing. Because the delay between nodes is allowed to change, the period of the pulse train must be adjusted as well. The information about these alterations is transmitted through the “slow” path. The “slow” clockwise path delay is altered by the same control mechanism as the fast path and exactly mirrors the “fast” path clockwise delay line, but it is not part of the control loop. By altering the “slow” clockwise path, the VCO can track the average delay around the system.

A further development, similar to the method proposed by B.K. Ahuja in (see Section 6.3), may include placing many sets of delay lines in parallel, each controlled by the same error signal. This would allow an entire bus structure to be constructed where the delays from node to node were forced to be the same.

One addition to the DSC structure can be made that includes the synchronization of the data path as well. Control signals to the pre and post delays within each node, as well as the propagation paths, can be replicated to form multiple sets of interconnections between adjacent nodes. If the delay between adjacent nodes can be equalized as well, the data will undergo the same delay between adjacent nodes allowing for a completely synchronized system. This proposal remains as future work for the author.

6.6) The distributed synchronous clock

The final step involved in the creation of a distributed synchronous clock required the combination of the counter-propagating multiple pulse generator circuit [figure 6-36] and the local node control circuit [figure 6-41]. This section will first provide an analytical proof that a steady-state solution exists for the distributed synchronous clock. Certain boundary conditions will be outlined and several different scenarios will be listed for comparison. Although only five nodes are assumed in this proof, the results can easily be generalized to any number of nodes, the algebraic equations are simply longer to derive. The second part of this section is a summary of an HSpice simulation of an eight-node distributed synchronous clock.

6.6.1) Analytical approach to the steady-state solution

An analytical model was created to determine if a steady-state solution to the distributed synchronous clocking strategy existed. The term 'solution' means that there is a particular reference frequency and a particular set of differential delay values, which allow for the synchronous arrival of pulses at every node. This proof does not consider how the system will arrive at the steady-state condition; it only shows that a steady-state solution exists. If the steady-state analytical proof were to indicate that there were either multiple solutions or no solution for a given set of parameters, then unstable behavior would be virtually inevitable.

The evidence that a steady-state solution exists is offered by way of example. The example considers a system with only 5 nodes. This example was chosen so that the mathematical equations were straight-forward enough to follow. Other examples have been carried out on systems with 6 and 7 nodes but are not included here, and the circuit simulation results presented in the next sections were done assuming 8 nodes. Although this does not comprise a true proof, it does, at minimum, lead to proof for a few specific cases. A true proof will be offered in future works by this author.

The following block diagram [figure 6-42] shows a simplified version of the distributed synchronous clock circuit of [figure 6-36]. In this example, there are five

nodes labeled A, B, C, D, and X. There are also five interconnection segments labeled L1, L2, L3, L4, and L5 that correspond to the delay between nodes. Each interconnection segment is made up of three identical transmission lines and each transmission line must have the same delay. However, each interconnection segment may differ in the amount of delay. The delays in both the clockwise and counter-clockwise paths are explicitly labeled in the diagram. The only delays that can be dynamically changed using the local control at a node, are the differential delays associated with each node. These delays are labeled ΔA , ΔB , ΔC , ΔD , and ΔX . A specific set of differential delays will exist for each scenario of the system.

For this example, there are exactly four independent, linear equations that can be formed. By equating the average clockwise delay from node "X" to node "N" (where "N" is either node A, B, C, or D) with the average counter-clockwise delay from node "X" to the node "N", a table of four equations can be made [Table 6-4]. The average delay is obtained by adding each incremental delay and dividing by the ideal total delay. A parameter called the ideal average delay "T" is introduced and is the ideal average delay between any two adjacent nodes. Although the actual value of "T" cancels in the algebra, the ratio does not. The value of "T" can then be calculated later as the exact period of pulse train required by the system.

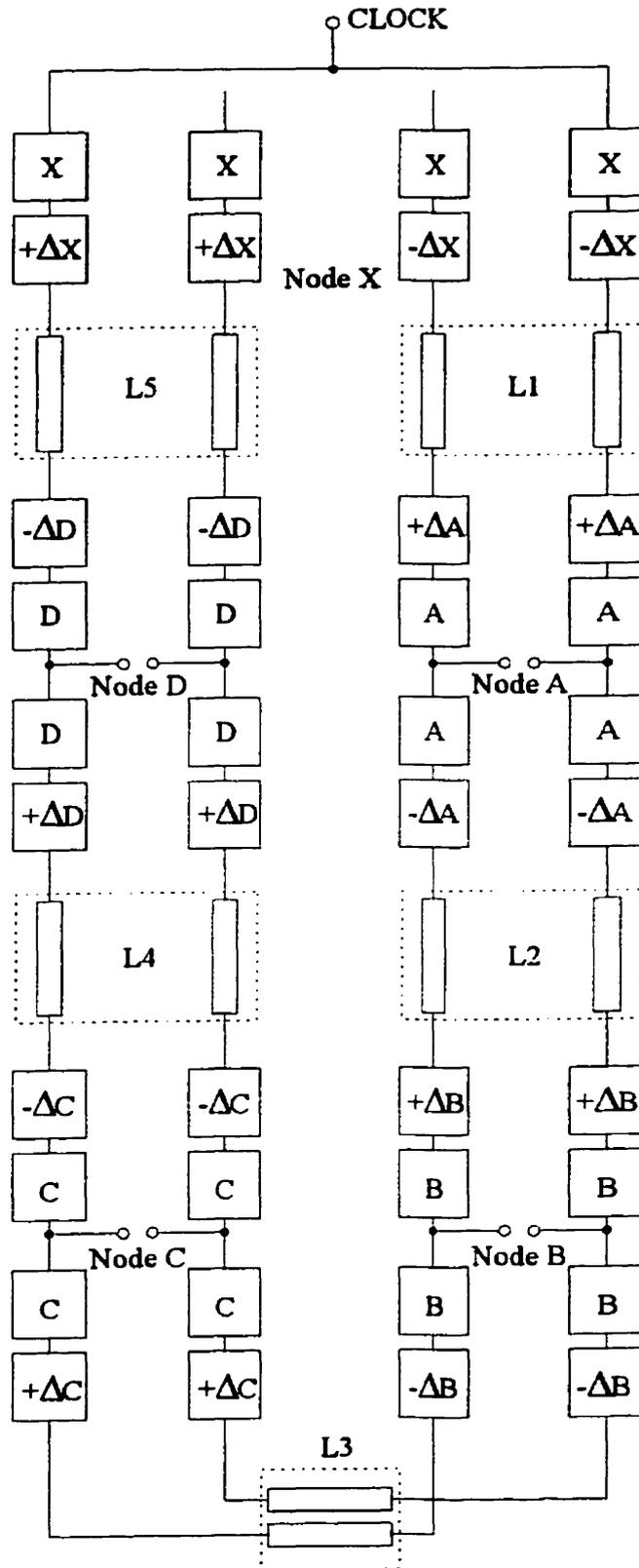


Figure 6-42: Analytical proof for block diagram of DSC

$$\begin{aligned}
1) \quad & \frac{X - \Delta X + L1 + \Delta A + A}{T} = \frac{X + \Delta X + L5 + 2D + L4 + 2C + L3 + 2B + L2 - \Delta A + A}{4T} \\
2) \quad & \frac{X - \Delta X + L1 + 2A + L2 + \Delta B + B}{2T} = \frac{X + \Delta X + L5 + 2D + L4 + 2C + L3 - \Delta B + B}{3T} \\
3) \quad & \frac{X - \Delta X + L1 + 2A + L2 + 2B + L3 + \Delta C + C}{3T} = \frac{X + \Delta X + L5 + 2D + L4 - \Delta C + C}{2T} \\
4) \quad & \frac{X - \Delta X + L1 + 2A + L2 + 2B + L3 + 2C + L4 + \Delta D + D}{4T} = \frac{X + \Delta X + L5 - \Delta D + D}{4T}
\end{aligned}$$

Table 6-4: Steady-state equations for model

The table [Table 6-5] describes several different scenarios for the fixed delays within the system (A, B, C, D, and X) as well as the delays for the transmission lines (L1, L2, L3, L4, and L5).

Delay	Scenario 1	Scenario 2	Scenario 3	Scenario 4
X	50	50	50	32
A	50	50	50	51
B	50	50	50	47
C	50	50	50	45
D	50	49	50	44
L1	50	50	50	53
L2	50	50	50	50
L3	50	50	50	61
L4	50	50	49	49
L5	50	50	50	30
ΔA	0	0.6	-0.2	2
ΔB	0	0.2	-0.4	-8
ΔC	0	-0.2	-0.6	-23
ΔD	0	-0.6	0.2	-23
T	150	149.6	149.8	138

Table 6-5: Different scenarios for the analytical proof

For any set of values for the fixed delays, there will always be a unique solution for ΔA , ΔB , ΔC , and ΔD . A table of values for several scenarios of fixed delay are given below, this includes the period of oscillation which is identical to the parameter "T".

It is crucial to understand the boundary conditions placed on this model. The basic mechanism used to alter the differential delays within a node is the PLL circuit. Hence,

the algorithm used to measure the difference in arriving pulses can produce only one compensating error signal. The differential delays within the node must all change in accordance with one error signal, thus, the magnitudes of all the changes to the differential delay elements must be the same (for example; node A has the variable delays: $+\Delta A$, $-\Delta A$, $+\Delta A$, and $-\Delta A$). Because the PLL algorithm is very limited, the delay of all three transmission lines within an interconnect segment must be equal. However, the delay of any of the interconnect segments, taken as a whole, can be different from one another. For example, L1 and L2 may be different, but the three lines within L1 must have identical delays. If the transmission lines within an interconnect segment are different, a steady-state value for the differential delays (ΔA , ΔB , ΔC , ΔD and ΔX) cannot be guaranteed. The PLL structures within each node may continually search for a steady state solution without reaching one, thus oscillatory behavior is possible. As discussed earlier, there are methods that allow for one transmission line between nodes, but for the sake of generality, this conditions has been outlined.

To allow for different transmission line delays within interconnect segments, a more complicated local control algorithm within each node would have to be used. The differential delays within the node must have different forward and reverse values. For example, node A must have two pairs of variable delays such as: $+\Delta A_1$, $-\Delta A_1$, $+\Delta A_2$, and $-\Delta A_2$). An algorithm that may allows for this would rely on a “learning” process such as a neural network. Neural networks are beyond the scope of this thesis and they will not be discussed further.

6.6.2) An HSpice simulation of the distributed synchronous clock

A complete HSpice simulation was built based on 1.2-micron MOSFET transistors. Most of the circuit was pure digital CMOS design, but several circuits involved analog designs using MOSFETS. The simulation is based on an eight-node synchronous distributed clock. The first node generates the pulses using what will be called the “Master-PLL” and the remaining 7 nodes help balance the system using what will be called the “Slave-PLLs”. The complete HSpice deck file can be found at the end of this chapter in APPENDIX B and is with reference to [figure 6-36] and [figure 6-41].

The simulation was run for 10000-nsec with 0.1-nsec step size. During the first 10-nsec of the simulation, all PLL action was disabled and the 3-bit counter was reset. During the next 100-nsec, all PLL action remained disabled, but the counter was allowed to increment. During this time, the VCO produced a nominal constant period pulse train of approximately 30-nsec that “loaded” each delay lines with the appropriate pulse train. After 110-nsec of simulation, the Master-PLL in node 1 was enabled and allowed to settle to a steady-state condition. This allowed the VCO to alter its period of oscillation slightly to match the total average delay of the system and to guarantee that there were 8 pulses in the “fast” delay lines. During the period from 5000-nsec to 10000-nsec, the Slave-PLLs within each node were enabled and the local control mechanism was allowed to function. The internal delays of each node were adjusted until the system settled producing eight synchronous waveforms at each of the nodes. The table [Table 6-6] is a summary of the simulation results obtained during the 10000-nsec simulation time. The figures [figure 6-43] and [figure 6-44] show the behavior prior to local-node control and after local node control.

Values of interconnect segments:	
L1	11.2 nsec
L2	9.8 nsec
L3	12.0 nsec
L4	11.0 nsec
L5	9.5 nsec
L6	10.0 nsec
L7	13.0 nsec
L8	10.7 nsec
Initial Period of Oscillation of the VCO:	30.48 nsec
Time Period between 0-nsec and 5000-nsec [figure 6-43]:	
Duration of Transient for Master-PLL:	2.648 μ sec
Steady-State value of analog voltage on VCO:	2.181 Volts
Period of oscillation of VCO in steady-state:	40.652 nsec
Total skew of rising-edge of node clock:	3.09 nsec
Total skew of falling-edge of node clock:	4.17 nsec
Time Period between 5000-nsec and 10000-nsec [figure 6-44]:	
Duration of Transient for Master-PLL:	4.1 μ sec
Steady-State value of analog voltage on VCO:	2.176 Volts
Period of oscillation of VCO in steady-state:	40.761 nsec
Total skew of rising-edge of node clock:	2.14 nsec
Total skew of falling-edge of node clock:	0.64 nsec

Table 6-6: Data collected from DSC HSpice simulation

A key feature highlighted by this table [Table 6-6] is that the distributed synchronous clock decreased the total skew of the falling-edges of all the clock signals at each node from 4.17-nsec to 0.64-nsec, an improvement of 73.4%. The fact that the falling-edge was brought into alignment is simply due to the sign of the PFD within each node, if the detection points were swapped and the error signals were swapped, then the rising-edge would be in alignment.

It is important to realize that this circuit simulation should be viewed primarily as a “proof-of-concept”. Aside from the result that the skew was decreased, this simulation indicates that the transient behavior is also stable. It shows that it is possible to achieve a steady-state solution using PLL concepts. There is much more work that can be carried out at this point. Faster PLL designs and a smaller MOSFET technology would be the first issues that could be addressed. Other implementations such as a completely analog approach using in-phase sinusoids and filters would also be very interesting to attempt.

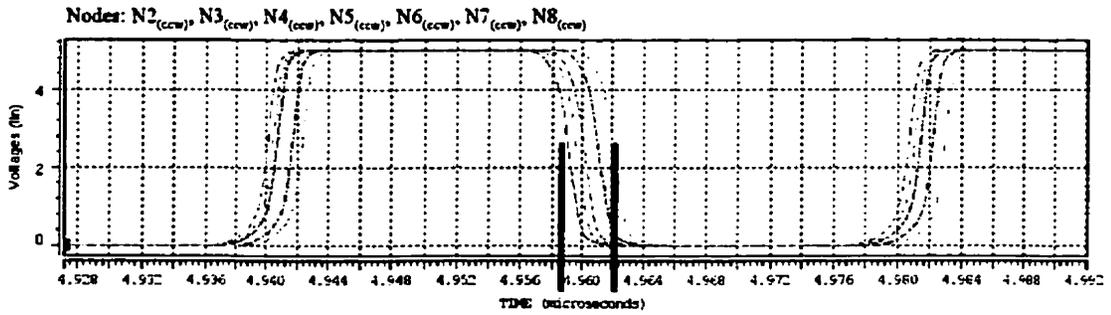
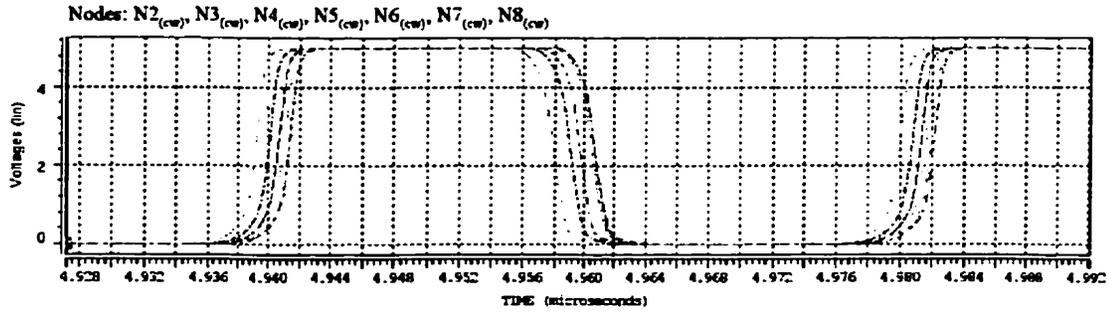


Figure 6-43: HSpice output of DSC system BEFORE local-control is activated

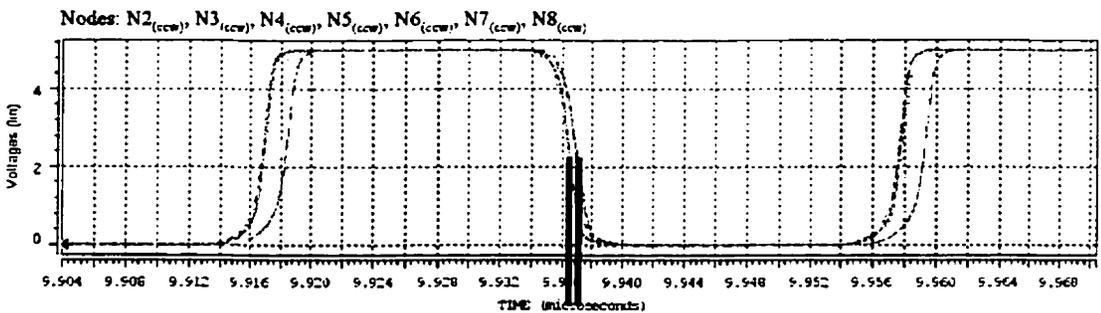
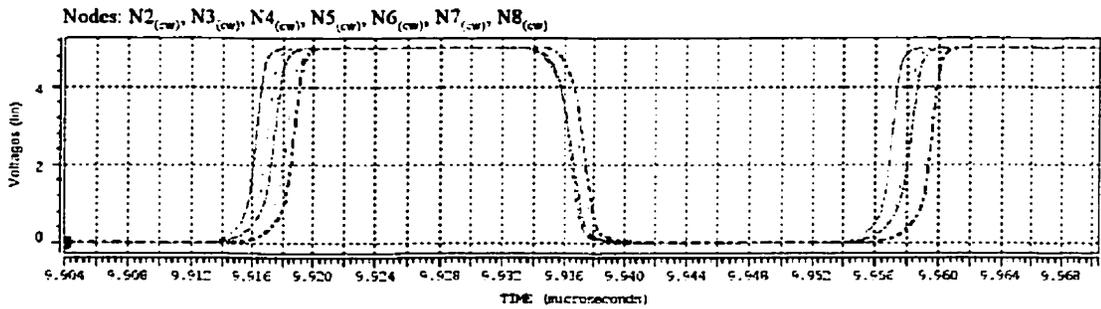


Figure 6-44: HSpice output of DSC system AFTER local-control is activated

6.7) .Conclusion

In this chapter, many standard synchronization techniques that allow high-speed computing systems to function were presented. The H-tree distribution network and the Phase-Locked Loop control system are two of the most commonly used techniques to provide synchronization. Many microprocessors possess both these mechanisms and significant chip area is devoted to their implementation. However, with systems that exceed a few square centimeters, these methods would be too costly and difficult to construct or perhaps even physically impossible due to the nature of the problem. Many methods have been employed in order to simplify and sometimes eliminate the requirements for synchronization, but these systems usually require more circuitry and protocols. Some methods simply lower the rate at which data is exchanged and other methods introduce complicated protocols and hardware buffering to avoid the problems of absolute synchronization. These methods usually cause latency and are more costly in terms of chip area because they require more transistors.

To achieve a fully integrated synchronous system connecting many spatially separated devices; a new method of clocking was developed called a distributed synchronous clock. The distributed synchronous clock is a technique that can achieve synchronization independently of the distance between adjacent nodes. Certain criteria were listed that must be followed for proper operation. However, the technique generally allows many nodes, which are separated by random amounts, to achieve and maintain synchronization using at most three connecting lines and PLL-type circuits. The use of PLLs within microprocessors is already very prevalent, and hence using PLLs in the distributed synchronous clock may be equally applicable. The distributed synchronous clock may in fact be able to link many microprocessors that are several centimeters away from each other.

Finally, the use of optical interconnects is also an important aspect of the distributed synchronous clocking system. Although this clocking technique can be implemented in electronics, the use of optics or optical fiber to distribute signals allows the clocking frequency to remain high, especially when long distances (several kilometers) are considered. Other implementations, such as satellite-to-satellite and satellite-to-ground synchronization may also be considered in future research.

6.8) References

- [1] V. Milutinovic, "Surviving the design of a 200 MHz RISC microprocessor : lessons learned", IEEE Computer Society Press, Los Alamos, Calif. 1997
- [2] P. Franzon, T. Schaffer, S. Lipa, A. Glaser, "Issues in chip-package codesign with MCM-D/flip-chip technology", IEEE Symposium on IC/Package Design Integration, 1998, pp. 88 -92
- [3] R.M. Reinschmidt, D.H. Leuthold, "Clocking considerations for a Pentium-based CPU module with 512K byte secondary cache", IEEE Multi-Chip Module Conference (MCMC-94), 1994 , pp. 26 -31
- [4] M. Nekili, G. Bois, Y. Savaria, "Pipelined H-trees for high-speed clocking of large integrated systems in presence of process variations", IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 5, No. 2, June 1997, pp. 161 -174
- [5] N.H.E. Weste, K. Eshraghian, Principles of CMOS VLSI Design 2nd Ed., Addison-Wessley, New York, 1992
- [6] S. Fischer, R. Senthinathan, H. Rangchi, H. Yardanmehr, "A 600 MHz IA-32 microprocessor with enhanced data streaming for graphics and video", IEEE International Solid-State Circuits Conference, Digest of Technical Papers, 1999, pp. 98 -101
- [7] E. Hirt, M. Scheffler, J.-P. Wyss, "Area I/O's potential for future processor systems", IEEE Micro Vol. 18, No. 4, July-Aug. 1998 , pp. 42 -49
- [8] UltraSPARC-IIi CPU Module, Sun Microsystems Data Sheets
- [9] A.M. Rincon, G. Cherichetti, J.A. Monzel, D.R. Stauffer, M.T. Trick, "Core design and system-on-a-chip integration" IEEE Design & Test of Computers, Vol. 14, No. 4, Oct.-Dec. 1997, pp. 26 -35
- [10] W.D. Tseng, K. Wang, "Fault coverage and defect level estimation models for partially testable MCMs", IEE Proceedings-Circuits, Devices and Systems, Vol. 147, No. 2, April 2000, pp. 119-24.
- [11] Altera Corporation < <http://www.altera.com/html/products/apex2.html#arch> >
- [12] Joseph Di Giacomo, Digital bus handbook, McGraw-Hill, New York, 1990.
- [13] International Technology Roadmap for Semiconductors <http://www.itrs.net/ntrs/publntrs.nsf>
- [14] Guillaume C. L. Boisset, Optomechanics and Optical Packaging for Free-Space Optical Interconnects, Ph.D. Thesis, McGill University, Montreal, Canada, 1998.
- [15] H. Willemsen, D. Nicholson, "GaAs ICs in commercial OC-192 equipment" 18th Annual Symposium on Gallium Arsenide Integrated Circuit (GaAs IC), 1996, pp. 10 -13
- [16] Cray Systems, <http://www.cray.com/products/systems/crayt3e> , 1999.
- [17] M. Muegge, M.G. Davis, "High performance bi-directional FIFO for bus interface applications", pp. 203-206
- [18] R. Vickers, "The development of ATM standards and technology: a retrospective", IEEE Micro, Vol. 13, No. 6, Dec. 1993 , pp. 62-73.
- [19] L. Hall, M. Clements, L. Wentai, G. Bilbro, "Clock distribution using cooperative ring oscillators", Symposium on VLSI Technology: Digest of Technical Papers, 1997, pp. 62-75.

- [20] W.D. Grover, "A new method for clock distribution", *IEEE Trans. On Circuits and Systems-I: Fundamental Theory and Applications*, Vol. 41, No. 2, Feb 1994, pp. 149-160
- [21] B.K. Ahuja, "Skew-free clock signal distribution network in a microprocessor", United States Patent, #5,307,381, Issued Apr. 26, 1994.
- [22] R.A. Omondi, The microarchitecture of pipelined and superscalar computers, Kluwer Academic Publishers, Boston, 1999.
- [23] R.A. Novotny, "Analysis of Smart Pixel Digital Logic and Optical Interconnects", Ph.D. Thesis, Heriot-Watt University, Edinburgh, Scotland, 1996.
- [24] S.M. Shinnars, Advanced modern control system theory and design, Wiley and Sons, New York, 1998
- [25] Ishibashi, K.; Goto, T.; Hayashi, T.; Okada, T.; Yamagiwa, A.; Shibata, M.; Akimoto, K.; Hamanaka, N.; Takahashi, T.; Koyama, A.; Aida, T., "Simultaneous Bidirectional Transceiver Logic", *IEEE Micro*, Vol. 19, No. 1, Jan.-Feb. 1999 , pp. 14 -19.
- [26] B. Razavi Ed., Monolithic Phase-Locked Loops and Phase Recovery Circuits, IEEE, New York, 1996.
- [27] F. Paganini, B. Bamieh, "Decentralization properties of optimal distributed controllers", *Proceedings of the 37th IEEE Conference on Decision and Control*, 1998. Vol. 2, 1998 , pp. 1877 -1882
- [28] R.E. Best, Phase-locked loops : design, simulation, and applications 4th Ed., McGraw-Hill, New York, 1999.
- [29] I.A. Young, J.K. Greason, K.L. Wong, "A PLL clock generator with 5 to 100 MHz of lock range for microprocessors", *IEEE Journal of Solid State Circuits*, Vol. 27, No. 11, Nov. 1992, pp. 1599-1606
- [30] A.B. Grebene, Bipolar and MOS analog integrated circuit design, J. Wiley and Sons, New York, 1984.

6.9) APPENDIX A

The Digital Phase-Lock Loop

There are many types of phase-lock loop (PLL) circuits ranging in speed and complexity and used for a variety of applications. The analog PLL is typically used as a lock-in circuit for television, cellular and microwave carriers, and most commonly used in FM radio carrier demodulation. The analog PLL often uses very sophisticated circuits involving frequency mixers, frequency doublers, and bandpass filters along with many other techniques, but they are usually suited for very high frequency, single frequency applications.

The type of PLL circuit explored in the following section is based on a lower frequency lock-in PLL required by some digital circuits. A digital PLL (DPLL) consists of the same basic theory as the analog PLL, however the building blocks are significantly more easy to design. This is not to infer that a DPLL is a trivial structure, only that the building blocks are more easily adapted to standard integrated circuit design. The following description of the DPLL is given because it is deemed an essential part of the distributed synchronous clock. The theory presented in this section is based on three sources. The textbook by E. Best [ref 28], a paper on the design of a digital phase-lock loop [ref 29] and numerous HSpice simulations.

The DPLL has four basic building blocks along with two additional building blocks that are optional depending on the application. The block diagram below [figure 6-A1] shows the negative feedback mechanism and the building blocks, with the non-

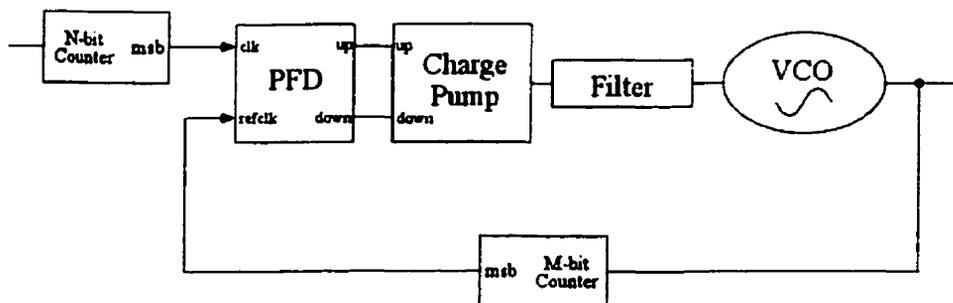


Figure 6-A1: A digital phase lock loop block diagram

essential ones darkened. This schematic demonstrates the feedback nature of the PLL and will be used later in this section to help build a linear control model of the DPLL.

The first element in the DPLL is the phase-frequency detector (PFD) which can detect mismatches of digital pulse trains in both phase and frequency. The PFD is essentially a 3-state finite-state machine that acts as the comparator between the reference signal and the feedback signal [figure 6-A2]. The PFD has two input signals and two output signals and is only triggered on

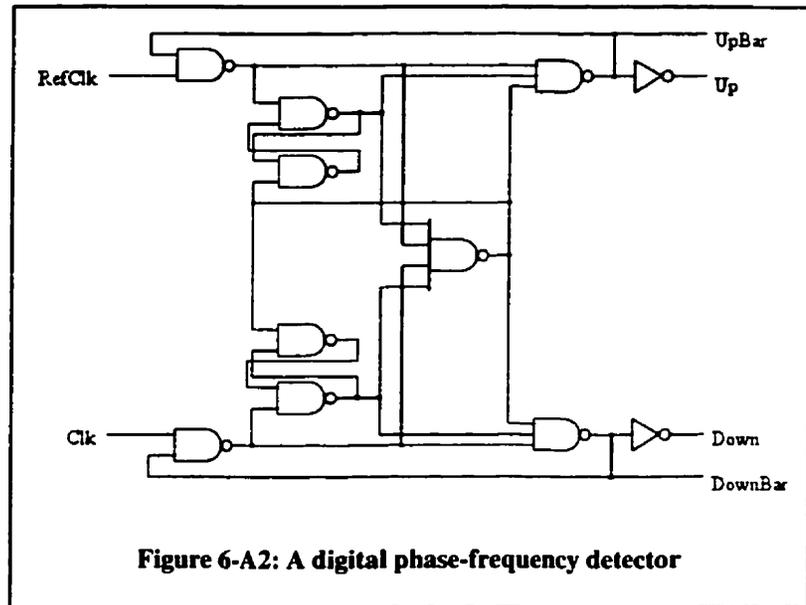


Figure 6-A2: A digital phase-frequency detector

the rising edges of the input signals, the waveforms are also given [figure 6-A3]. The PFD encodes the error signal in two ways; the sign of the error is encoded by choosing the 'up' or 'down' output and the magnitude of the error is encoded with the duration of the pulse. The linear, small-signal model for the PFD is [Eqn. 6-A1], where Θ_i and Θ_o are the phases of the input signals, ' K_c ' is the dc-gain (and also the slope of the curve), and v_e is the average voltage error. A plot of average output voltage versus the input phase difference including is given below [figure 6-A4].

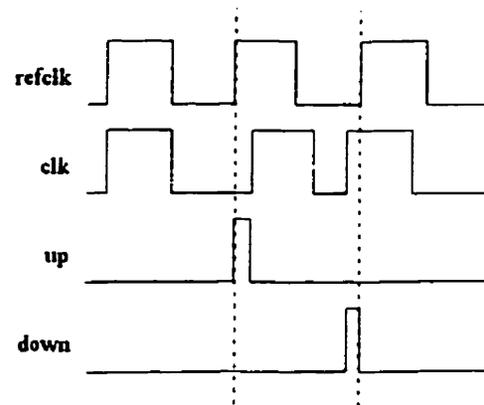


Figure 6-A3: PFD waveforms

Eqn. 6-A1:

$$v_e = K_c(\Theta_i - \Theta_o)$$

The second element of the DPLL is the charge pump. The charge pump is the first part of a very simple digital to analog (D/A) converter where the digital voltage pulses are converted to current pulses.

The duration of the pulses is still preserved where the sign of the error is indicated by the direction of current flow. The charge pump also has a dc-gain and is equal to the magnitude of the current pulse divided by the magnitude of the voltage pulse. One unique quality of the charge pump is that during the interval when both the outputs of the PFD are low, the output voltage of the charge pump remains at the previously adjusted level; it has the ability to hold the voltage at the most recently adjusted voltage level. A typical charge pump is shown here [figure 6-A5] where the ratio of input voltage to

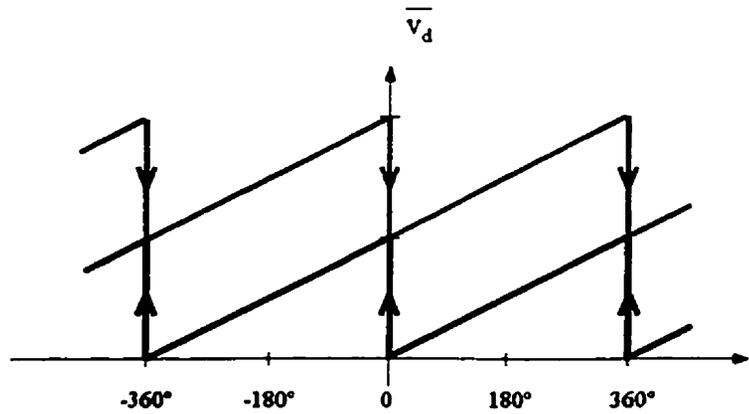


Figure 6-A4: PDF transfer function

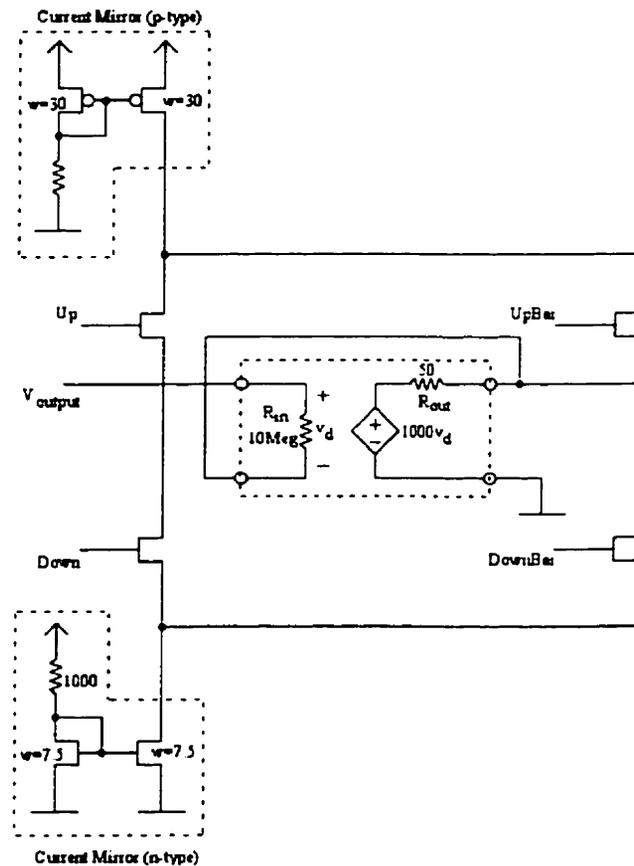


Figure 6-A5: A charge-pump

output current can be calculated depending on the magnitude of the currents produced by the current mirrors. To obtain the nominal output current, a +2.5-volt load is placed at the output such that current can either be pushed or pulled into or out of the voltage source. The linear, small-signal model for the charge pump is: $i_p = K_p \cdot v_e$, where 'K_p' is the dc-gain, and i_p is the current pulse.

The filter is the third essential part of the PLL and is also the second part of the D/A converter involving the charge pump. The filter will present a quasi-dc control voltage to the VCO, which is derived from the PFD error signal. The filter is also used to

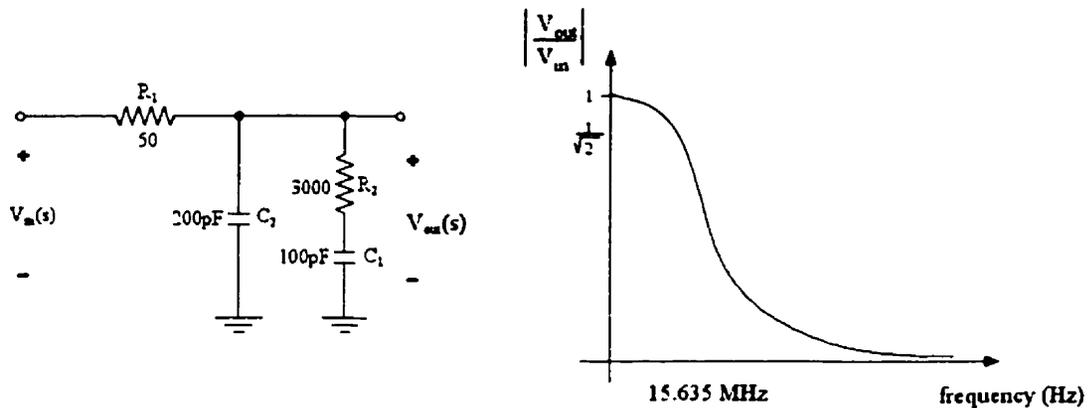


Figure 6-A6: Two-pole low-pass filter

eliminate high frequency noise from the error signal while at the same time converting the current pulses from the charge pump into a piece-wise stepped voltage for the VCO. The filter is shown below as well as the magnitude of the frequency response for two different values of capacitor 'C₂' [figure 6-A6]. The filter has two poles due to both capacitors and is used in order to make the response slightly steeper at the -3dB point. If the bandwidth is made very large, then the settling time will be very short, however, the PLL is then susceptible to unwanted noise components in the system such as power supply noise and can cause the VCO frequency output to jitter. The capacitor 'C₂' in the filter provides an additional feature; the capacitor helps hold the voltage at the input of the VCO during periods when the charge pump is not producing current pulses. For example, if the error signal causes the voltage on C₂ to rise from 2.5-volts to 3.1-volts

and then stabilize, the VCO will increase in frequency and then remain at the higher frequency when the derivative of the error has gone to zero. The voltage-to-current transfer function is given below for this filter where the pole at zero frequency can be easily seen (i.e.: with zero input, the 1/s integrator produces a constant, non-zero, output).

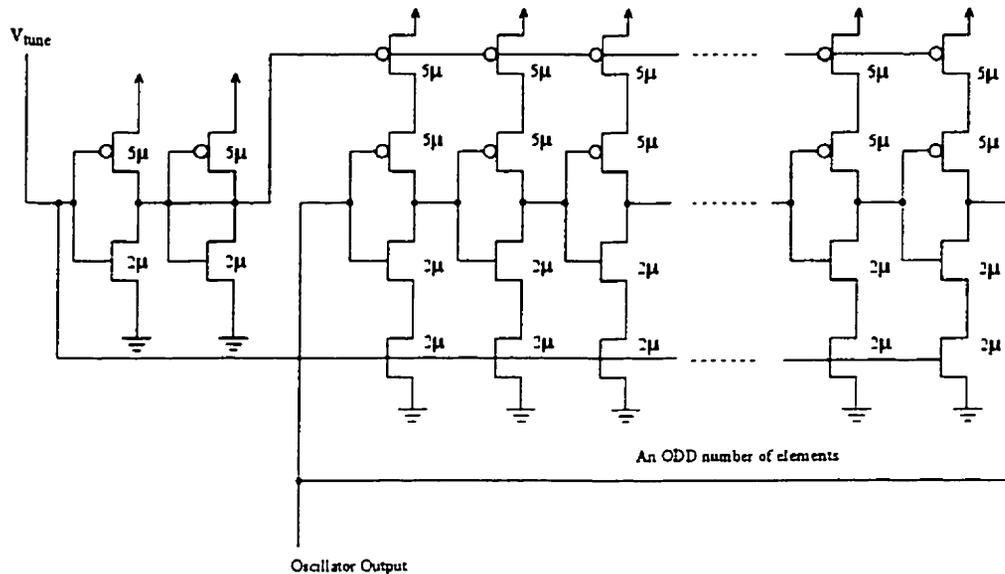


Figure 6-A7: A CMOS voltage controlled oscillator

Eqn. 6-A2:
$$G(s) = \frac{sR_2C_1 + 1}{s^2C_1R_2C_2 + s(C_1 + C_2)}$$

The final element of the PLL is the VCO. The following is a circuit of a CMOS VCO [figure 6-A7] and a plot of control voltage versus angular frequency is also given [figure 6-A8]. The VCO circuit is based on current-staved CMOS inverters that act as variable delay elements. The current-staved inverter uses an additional pMOS and nMOS transistor in the path from Power and Ground, respectively. As the bias on the gates of these inverters change, so does the effective resistance, thereby effecting the charge and discharge of the output. For this particular VCO, the center frequency is 58.5-MHz at a 2.5-volt bias and its linear range lies between approximately 40-MHz and 100-MHz. By integrating the output voltage of the VCO [Eqn. 6-A2], the linear small-signal model of

the VCO can be obtained [Eqn. 6-A3] where the constant 'K_{vco}' is given by the slope of the line in the plot, its units are radians/sec/volt. The unknown in absolute phase is obvious after the integral has been taken.

However, for the purposes of using this model in a control systems analysis, the Laplace transform is more appropriate, the constant is ignored just as the forcing function is ignored to determine the homogeneous solution of a differential equation.

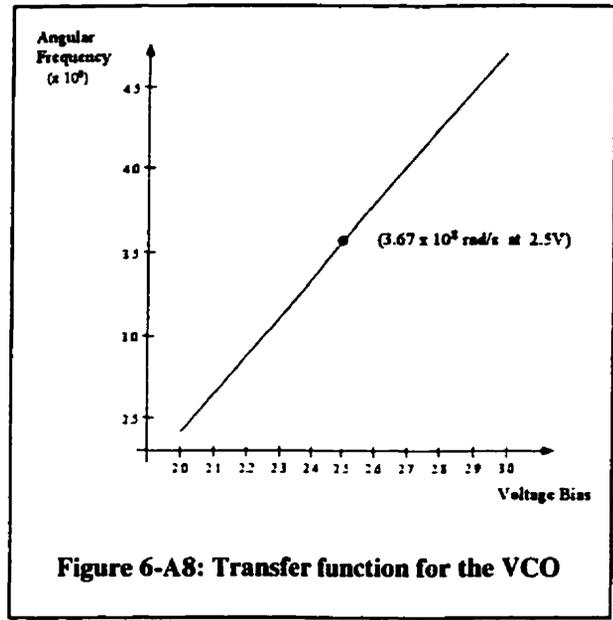


Figure 6-A8: Transfer function for the VCO

$$\theta(t) = \int K_{vco} \cdot v_{dc}(t) \cdot dt$$

$$\theta(t) = K_{vco} \int v_{dc}(t) \cdot dt + C$$

Eqn. 6-A3 a,b,c)

$$\Theta(s) = \frac{K_{vco} \cdot V_{dc}(s)}{s}$$

These four elements; the PFD, the charge pump, the filter, and the VCO are then connected in a negative feedback configuration as shown earlier [figure 6-A1]. The results of an HSpice simulation are given here [figure 6-A9] and show the transient response of the control signal into the VCO. In the simulation, the input reference clock was 57.2-MHz and at time = 4-ms, a unit-step phase error of roughly 45° is introduced into the reference clock. The filter is a wide band filter with C₂ = 20-pF. The more narrow the band, the slower the PLL reacts because less energy is allowed to pass through the filter, however, this means that less noise energy is also allow through the filter.

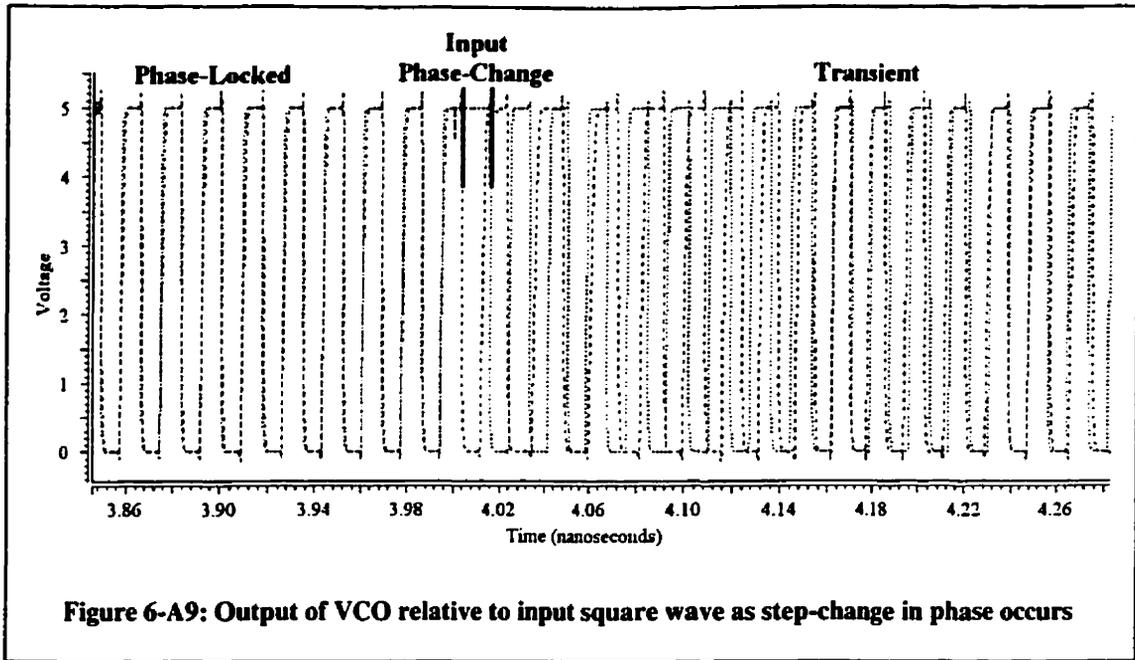


Figure 6-A9: Output of VCO relative to input square wave as step-change in phase occurs

To verify the HSpice simulations and apply a little theory to the PLL results, a first-order linear feedback system was modeled using the small-signal equations developed from each sub-circuit. The list of equations for the sub-circuits is re-written here:

$$\begin{aligned} \text{PFD:} \quad & V_e(s) = K_e (\Theta_i(s) - \Theta_o(s)) \\ \text{Charge Pump:} \quad & I_p(s) = K_p V_e(s) \\ \text{Filter:} \quad & V_c(s) = \frac{sR_2C_1 + 1}{s^2C_1R_2C_2 + s(C_1 + C_2)} I_p(s) \\ \text{VCO:} \quad & \Theta_o(s) = \frac{K_{vco}}{s} V_c(s) \end{aligned}$$

Using the feedback model shown above [figure 6-A1], the overall transfer function relating the voltage $V_c(s)$ to the phase input $\Theta_i(s)$ can be derived [Eqn. 6-A4]:

$$\text{Eqn. 6-A4} \quad V_c(s) = \frac{R_2C_1K_pK_e s^2 + K_pK_e s^1}{C_1R_2C_2s^3 + (C_1 + C_2)s^2 + R_2C_1K_pK_eK_{vco}s + K_pK_eK_{vco}} \Theta_i(s)$$

where:

$$R_1 = 50\text{-}\Omega$$

$$R_2 = 3000\text{-}\Omega$$

$$C_1 = 100\text{-pF}$$

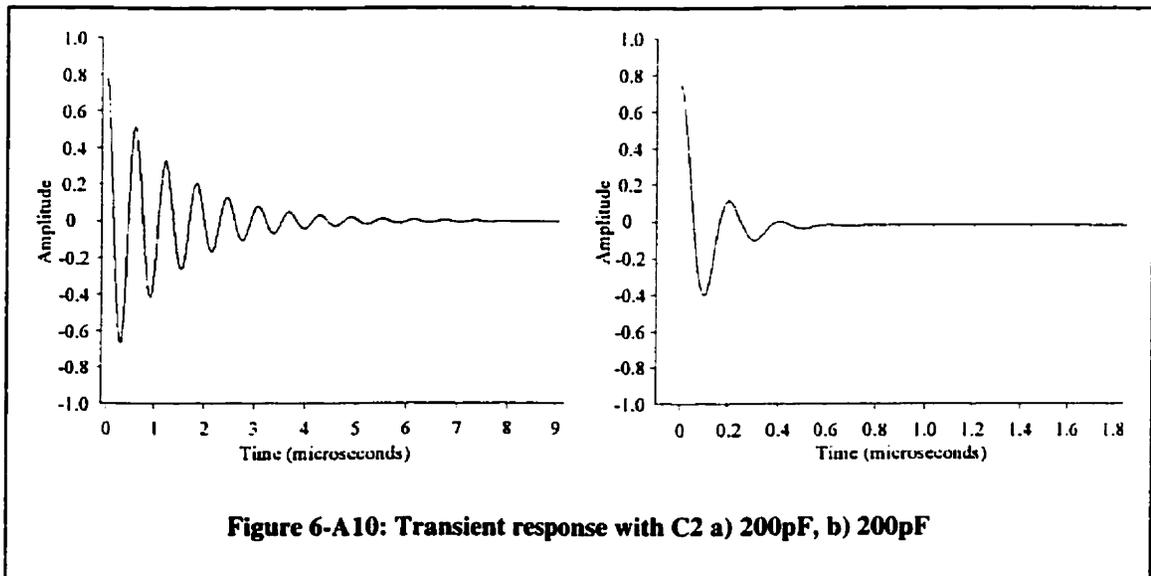
$$C_2 = 200\text{-pF and } 20\text{-pF}$$

$$K_c = 0.786$$

$$K_p = 0.00012$$

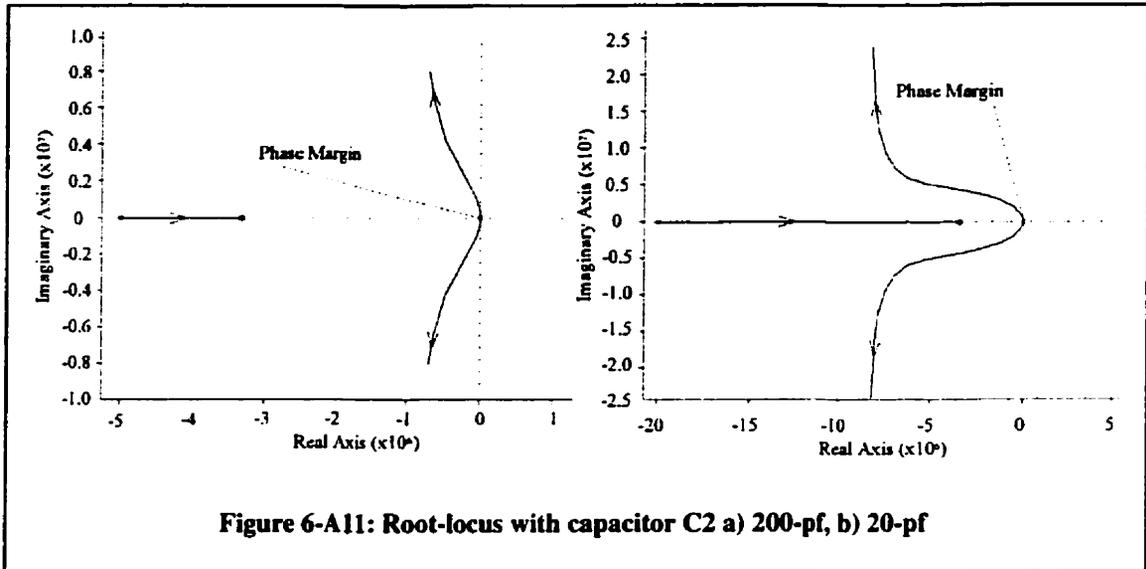
$$K_{vco} = 2.3503e8$$

The transient response of this model was done using MATLAB by applying a step response to the transfer function. The responses for the two values of 'C₂' are given below [figure 6-A10]. The method of root-locus was also done using the open-loop transfer function of the PLL system [figure 6-A11]. The root locus method creates its own closed loop system and varies the feedback gain from 1 to infinity, it then traces out the positions of the effective poles and zeros of the system. A standard method used by



designers is to ensure that the poles in the right half plane of the pole-zero plot are at an angle of less than 45° from the negative 'x' axis. This allows for more stable operation of the circuit [ref 30] by damping the system quickly. The capacitor value of 20-pF for C₂ ensures that for the dc-gain of the PLL, the damping is sufficiently large. As for the value of 200-pF for C₂, damping still occurs, but no matter what dc-gain is available, the poles

of the system never lie more than 45° from the negative x-axis; called the 'Phase Margin'.



The final treatment of the DPLL is its ability to do frequency conversion. The two darkened elements shown in the block diagram above [figure 6-A1] are synchronous counters. The counter at the input is an 'M' bit counter, and the counter in the feedback loop in an 'N' bit counter. These counters are typically placed internal to the chip so that they can operate very high speeds. A low frequency input square wave ' f_i ' can be frequency converted up to a very high internal clock rate ' f_o ' by the factor: $f_o = (N/M) f_i$. One condition, however, is that the VCO must be able to operate at a frequency of $(N/M) f_i$ for any value of M and N chosen. The PLL frequency conversion works by dividing the frequency of the output of the VCO by 'N' and comparing this with the lower frequency $(1/M)f_i$. Since there must be a condition for both inputs of the PFD to produce a zero error, the frequency of the VCO will be adjusted to whatever value is required in order to produce the proper frequency out of the 'N' bit counter.

6.10) APPENDIX B

The following HSpice circuit description is a transistor-level description of the multiple event generator [figure 6-36] and the local control node [figure 6-41] combined into one complete circuit.

```
* A Distributed Synchronous Clocking
* distribution system
* INCLUDING Local Mode PLLs
*
* Includes a 3-bit synchronous counter
* A 7-node clockwise slow delay line
* A 7-node clockwise fast delay line
* A 7-node counter-clockwise fast delay line
* - each include ~10 nsec long transmission lines
*
*
*.....
*..... Models .....
*
.model pmosfet PMOS
*   wmin=2.0e-06   wmax=500e-06
+   level=3       vto=-.84
+   gamma=.53     phi=.58
+   rd=94         rs=94
+   is=1e-16      pb=.8
+   cgso=3.284e-10  cgdo=3.284e-10
+   rsh=100       cj=0.00041
+   mj=.54        cjsw=3.4e-10
+   mjsw=.3       js=0.0001
+   tox=2.5e-08   nsub=1.75e+16
+   nfs=8.4e+11  tpg=1
+   xj=0          ld=6e-08
+   uo=205        vmax=500000
+   fc=.5         delta=.4598
+   theta=.14     eta=.17
+   kappa=10
*
.model nmosfet NMOS
*   wmin=2.0e-06   wmax=500e-06
+   level=3       vto=.79
+   gamma=.38     phi=.53
+   rd=63         rs=63
+   is=1e-16      pb=.8
+   cgso=1.973e-10  cgdo=1.973e-10
+   rsh=45        cj=0.00029
+   mj=.486       cjsw=3.3e-10
+   mjsw=.33      js=0.0001
+   tox=2.5e-08   nsub=8.7e+15
+   nfs=8.2e+11  tpg=1
+   xj=1e-07      ld=7e-08
+   uo=577        vmax=150000
+   fc=.5         delta=.3551
+   theta=0.046   eta=.16
+   kappa=0.05
*
*.....
*..... Inverter .....
*
.subckt inv 1 2 99
*
* A -----| |
* Out -----| |
* Vdd -----| |

```

```

*
mp1 2 1 99 99 pmosfet l=1.2u w=5u
mn1 2 1 0 0 nmosfet l=1.2u w=2u
*
.ends inv
*
***** NAND 2 *****
*
.subckt nand2 1 2 3 99
*
* A -----| | |
* B -----| | |
* Out -----|
* Vdd
*
mp1 3 1 99 99 pmosfet l=1.2u w=5u
mp2 3 2 99 99 pmosfet l=1.2u w=5u
mn1 4 1 0 0 nmosfet l=1.2u w=2u
mn2 3 2 4 0 nmosfet l=1.2u w=2u
*
.ends nand2
*
***** NAND 3 *****
*
.subckt nand3 1 2 3 4 99
*
* A -----| | | |
* B -----| | | |
* C -----| | | |
* Out -----|
* Vdd
*
mp1 4 1 99 99 pmosfet l=1.2u w=5u
mp2 4 2 99 99 pmosfet l=1.2u w=5u
mp3 4 3 99 99 pmosfet l=1.2u w=5u
mn1 6 1 0 0 nmosfet l=1.2u w=2u
mn2 5 2 6 0 nmosfet l=1.2u w=2u
mn3 4 3 5 0 nmosfet l=1.2u w=2u
*
.ends nand3
*
***** NAND 4 *****
*
.subckt nand4 1 2 3 4 5 99
*
* A -----| | | | |
* B -----| | | | |
* C -----| | | | |
* D -----| | | | |
* Out -----|
* Vdd
*
mp1 5 1 99 99 pmosfet l=1.2u w=5u
mp2 5 2 99 99 pmosfet l=1.2u w=5u
mp3 5 3 99 99 pmosfet l=1.2u w=5u
mp4 5 4 99 99 pmosfet l=1.2u w=5u
mn1 8 1 0 0 nmosfet l=1.2u w=2u
mn2 7 2 8 0 nmosfet l=1.2u w=2u
mn3 6 3 7 0 nmosfet l=1.2u w=2u
mn4 5 4 6 0 nmosfet l=1.2u w=2u
*
.ends nand4
*
***** PFD *****
*
.subckt pfd 1 2 3 5 99
*
| | | |

```

```

* Clk      ---| | | |
* RefClk   -----| | |
* UpBar    -----| | |
* DownBar  -----| | |
* Vdd
*
Xgate1     1 3 8 99 nand2
Xgate2     8 9 11 99 nand2
Xgate3     11 13 9 99 nand2
Xgate4     13 12 10 99 nand2
Xgate5     10 7 12 99 nand2
Xgate6     2 5 7 99 nand2
Xgate7     8 11 13 3 99 nand3
Xgate8     13 12 7 5 99 nand3
Xgate9     11 8 7 12 13 99 nand4
Xgate10    3 4 99 inv
Xgate11    5 6 99 inv
*
.ends pfd
*
*****
***** P-Mirror *****
*
.subckt pmirror 1 99
*
* OutCurrent ---|
* Vdd
*
Mp1 2 2 99 99 pmosfet l=1.2u w=10u
Mp2 1 2 99 99 pmosfet l=1.2u w=10u
Rload 2 0 1600
*
.ends pmirror
*
*****
***** N-Mirror *****
*
.subckt nmirror 1 99
*
* OutCurrent ---|
* Vdd
*
Mn1 2 2 0 0 nmosfet l=1.2u w=10u
Mn2 1 2 0 0 nmosfet l=1.2u w=10u
Rload 99 2 4400
*
.ends nmirror
*
*****
***** Charge Pump *****
*
.subckt chargepump 1 2 3 4 5 99
*
* upbar  -----| | | |
* up     -----| | | |
* downbar -----| | | |
* down   -----| | | |
* output -----| | | |
* Vdd
*
Xupbias    7 99 pmirror
Xdnbias    8 99 nmirror
*
Mup        7 2 5 0 nmosfet l=1.2u w=15u
Mdown      5 4 8 0 nmosfet l=1.2u w=15u
Mupbar     7 1 6 0 nmosfet l=1.2u w=15u
Mdnbar     6 3 8 0 nmosfet l=1.2u w=15u
*
Eamp       9 0 5 6 1000
Rin_amp    5 6 10Meg
Rout_amp   9 6 50

```

```

*
.ends chargepump
*
*****
***** RC filter *****
*
.subckt filter      1 2
*
* Input Voltage  ---| |
* Output Voltage -----|
*
R1  1 2 50
R2  2 3 3k
C1  3 0 100pf
C2  2 0 200pf
*
* Note: C2 = 200pf means slow transient decay because
* the filter is a narrow low pass filter and less noise
* is passed by the filter. This capacitor is also used
* because the chain takes a long time to react to changes
* and thus only small perturbations can be accomadated.
*
.ends filter
*
*****
***** Delay Element *****
*
.subckt delayel    1 2 3 4 99
*
* In      -----| | | |
* VcN     -----| | | |
* VcP     -----| | | |
* Out     -----| | | |
* Vdd
*
mp1 6 3 99 99 pmosfet l=1.2u w=5u
mp2 4 1 6 99 pmosfet l=1.2u w=5u
mn2 4 1 7 0 nmosfet l=1.2u w=2u
mn1 7 2 0 0 nmosfet l=1.2u w=2u
*
.ends delayel
*
*****
***** VCO *****
*
.subckt vco      98 1 99
*
* Bias      ---| |
* FrequencyOut -----|
* Vdd
*
Vconstantdc      97 96 dc 5V
Ecomplement      96 0 98 0 -1
*
Xdelayel1      1 98 97 2 99 delayel
Xdelayel2      2 98 97 3 99 delayel
Xdelayel3      3 98 97 4 99 delayel
Xdelayel4      4 98 97 5 99 delayel
Xdelayel5      5 98 97 6 99 delayel
Xdelayel6      6 98 97 7 99 delayel
Xdelayel7      7 98 97 8 99 delayel
Xdelayel8      8 98 97 9 99 delayel
Xdelayel9      9 98 97 10 99 delayel
Xdelayel10     10 98 97 11 99 delayel
Xdelayel11     11 98 97 12 99 delayel
Xdelayel12     12 98 97 13 99 delayel
Xdelayel13     13 98 97 14 99 delayel
Xdelayel14     14 98 97 15 99 delayel
Xdelayel15     15 98 97 16 99 delayel
Xdelayel16     16 98 97 17 99 delayel
Xdelayel17     17 98 97 18 99 delayel

```

```

Xdelay18 18 98 97 19 99 delayel
Xdelay19 19 98 97 20 99 delayel
Xdelay20 20 98 97 21 99 delayel
Xdelay21 21 98 97 22 99 delayel
Xdelay22 22 98 97 23 99 delayel
Xdelay23 23 98 97 24 99 delayel
Xdelay24 24 98 97 25 99 delayel
Xdelay25 25 98 97 26 99 delayel
Xdelay26 26 98 97 27 99 delayel
Xdelay27 27 98 97 28 99 delayel
Xdelay28 28 98 97 29 99 delayel
Xdelay29 29 98 97 30 99 delayel
Xdelay30 30 98 97 31 99 delayel
Xdelay31 31 98 97 32 99 delayel
Xdelay32 32 98 97 33 99 delayel
Xdelay33 33 98 97 34 99 delayel
Xdelay34 34 98 97 35 99 delayel
Xdelay35 35 98 97 36 99 delayel
Xdelay36 36 98 97 37 99 delayel
Xdelay37 37 98 97 1 99 delayel
*
* Period of this VCO is 30 nsec
*
*
.ends vco
*
*****
***** Controller *****
*
.subckt controller 1 2 4 3 6 5 7 99
*
* UpBarIn -----| | | | |
* DownBarIn -----| | | | |
* UpBarOut -----| | | | |
* UpOut -----| | | | |
* DownBarOut -----| | | | |
* DownOut -----| | | | |
* control -----| | | | |
* Vdd
*
Xnand21 1 7 3 99 nand2
Xnand22 2 7 5 99 nand2
Xinv1 3 4 99 inv
Xinv2 5 6 99 inv
*
.ends controller
*
*****
***** CMOS D-Filp Flop *****
*
.subckt dffreset 1 5 2 98 99
*
* Input -----| | | |
* Output -----| | | |
* Clock -----| | | |
* Reset -----| | | |
* Vdd
*
mpclk 3 2 99 99 pmosfet l=1.2u w=4u
mnclk 3 2 0 0 nmosfet l=1.2u w=2u
mpreset 4 98 99 99 pmosfet l=1.2u w=4u
mnreset 4 98 0 0 nmosfet l=1.2u w=2u
*
mp1 10 1 99 99 pmosfet l=1.2u w=4u
mp2 7 2 10 99 pmosfet l=1.2u w=4u
mp3 8 7 99 99 pmosfet l=1.2u w=4u
mp4 9 4 99 99 pmosfet l=1.2u w=4u
mp5 9 8 99 99 pmosfet l=1.2u w=4u
mp6 7 3 9 99 pmosfet l=1.2u w=4u
mp7 8 3 11 99 pmosfet l=1.2u w=4u
mp8 12 14 99 99 pmosfet l=1.2u w=4u

```

```

mp9      11  2 12 99 pmosfet l=1.2u w=4u
mp10    14 11 99 99 pmosfet l=1.2u w=4u
mp11    14  4 99 99 pmosfet l=1.2u w=4u
mp12     5 14 99 99 pmosfet l=1.2u w=4u
*
mn1      6  1  0  0 nmosfet l=1.2u w=2u
mn2      7  3  6  0 nmosfet l=1.2u w=2u
mn3      8  7  0  0 nmosfet l=1.2u w=2u
mn4     15  4  0  0 nmosfet l=1.2u w=2u
mn5     16  8 15  0 nmosfet l=1.2u w=2u
mn6      7  2 16  0 nmosfet l=1.2u w=2u
mn7      8  2 11  0 nmosfet l=1.2u w=2u
mn8     13 14  0  0 nmosfet l=1.2u w=2u
mn9     11  3 13  0 nmosfet l=1.2u w=2u
mn10    14 11 17  0 nmosfet l=1.2u w=2u
mn11    17  4  0  0 nmosfet l=1.2u w=2u
mn12     5 14  0  0 nmosfet l=1.2u w=2u
*
.ends dffreset
*
*****
***** XNOR2 *****
*
.subckt xnor2 1 3 8 99
*
* In1 -----| | |
* In2 -----| | |
* Out -----| | |
* Vdd
*
mp1  2  1 99 99 pmosfet l=1.2u w=5u
mp2  4  3 99 99 pmosfet l=1.2u w=5u
mp3  5  2 99 99 pmosfet l=1.2u w=5u
mp4  5  3 99 99 pmosfet l=1.2u w=5u
mp5  8  1  5 99 pmosfet l=1.2u w=5u
mp6  8  4  5 99 pmosfet l=1.2u w=5u
mn1  2  1  0  0 nmosfet l=1.2u w=2u
mn2  4  3  0  0 nmosfet l=1.2u w=2u
mn3  8  1  6  0 nmosfet l=1.2u w=2u
mn4  6  4  0  0 nmosfet l=1.2u w=2u
mn5  8  2  7  0 nmosfet l=1.2u w=2u
mn6  7  3  0  0 nmosfet l=1.2u w=2u
*
.ends xnor2
*
*****
***** 4-bit counter *****
*
.subckt counter4bit 98 97 1 2 3 4 99
*
* clock -----| | | | |
* reset -----| | | | |
* Bit1 -----| | | | |
* Bit2 -----| | | | |
* Bit3 -----| | | | |
* Bit4 -----| | | | |
*
* reset is active HIGH (if it is high, the counter resets)
*
XxnorA    0  1  5 99 xnor2
Xgffa    5  1 98 97 99 dffreset
XinvB     1  6 99 inv
XxnorB    6  2  7 99 xnor2
Xdffb    7  2 98 97 99 dffreset
XnandC    1  2  8 99 nand2
XxnorC    8  3  9 99 xnor2
Xgffc    9  3 98 97 99 dffreset
XnandD    1  2  3 10 99 nand3
XxnorD   10  4 11 99 xnor2
Xdffd   11  4 98 97 99 dffreset
*

```

```

.ends counter4bit
*
*****
***** Capacitor Delay Line *****
*
.subckt capdelayline 1 98 8 97 15 99
*
* In -----| | | |
* PreCapBias -----| | | |
* Observe -----| | | |
* PostCapBias -----| | | |
* Out -----| | | |
* Vdd
*
Xdelay1      1 2 99 inv
Xdelay2      2 3 99 inv
Xdelay3      3 98 97 4 99 delayel
M3           0 4 0 98 pmosfet l=1.2u w=100u
Xdelay4      4 98 97 5 99 delayel
M4           0 5 0 98 pmosfet l=1.2u w=100u
Xdelay5      5 98 97 6 99 delayel
M5           0 6 0 98 pmosfet l=1.2u w=100u
Xdelay6      6 98 97 7 99 delayel
M6           0 7 0 98 pmosfet l=1.2u w=100u
Xdelay7      7 8 99 inv
Xdelay8      8 9 99 inv
Xdelay9      9 97 98 10 99 delayel
M9           0 10 0 97 pmosfet l=1.2u w=100u
Xdelay10     10 97 98 11 99 delayel
M10          0 11 0 97 pmosfet l=1.2u w=100u
Xdelay11     11 97 98 12 99 delayel
M11          0 12 0 98 pmosfet l=1.2u w=100u
Xdelay12     12 97 98 13 99 delayel
M12          0 13 0 98 pmosfet l=1.2u w=100u
Xdelay13     13 14 99 inv
Xdelay14     14 15 99 inv
*
.ends capdelayline
*
*****
***** Fixed Delay *****
*
.subckt fixdelay 1 9 99
*
* In -----| |
* Out -----| |
* Vdd
*
Vbias        98 0 dc 2.5V
*
Xdelay1      1 2 99 inv
Xdelay2      2 3 99 inv
Xdelay3      3 98 98 4 99 delayel
M3           0 4 0 98 pmosfet l=1.2u w=100u
Xdelay4      4 98 98 5 99 delayel
M4           0 5 0 98 pmosfet l=1.2u w=100u
Xdelay5      5 98 98 6 99 delayel
M5           0 6 0 98 pmosfet l=1.2u w=100u
Xdelay6      6 98 98 7 99 delayel
M6           0 7 0 98 pmosfet l=1.2u w=100u
Xdelay7      7 8 99 inv
Xdelay8      8 9 99 inv
*
.ends fixdelay
*
*****
***** Transmission Lines *****
*
.subckt tlineL1 1 4
*
* In -----| |

```

```

* Out -----|
*
Esource      2 0 1 0 1
Tline1      2 0 3 0 Z0=50 TD=11.2n
Rload1      3 0 50
Eload2      4 0 3 0 1
*
.ends tlineL1
*
.subckt tlineL2 1 4
*
* In -----|
* Out -----|
*
Esource      2 0 1 0 1
Tline1      2 0 3 0 Z0=50 TD=9.8n
Rload1      3 0 50
Eload2      4 0 3 0 1
*
.ends tlineL2
*
.subckt tlineL3 1 4
*
* In -----|
* Out -----|
*
Esource      2 0 1 0 1
Tline1      2 0 3 0 Z0=50 TD=12n
Rload1      3 0 50
Eload2      4 0 3 0 1
*
.ends tlineL3
*
.subckt tlineL4 1 4
*
* In -----|
* Out -----|
*
Esource      2 0 1 0 1
Tline1      2 0 3 0 Z0=50 TD=11n
Rload1      3 0 50
Eload2      4 0 3 0 1
*
.ends tlineL4
*
.subckt tlineL5 1 4
*
* In -----|
* Out -----|
*
Esource      2 0 1 0 1
Tline1      2 0 3 0 Z0=50 TD=9.5n
Rload1      3 0 50
Eload2      4 0 3 0 1
*
.ends tlineL5
*
.subckt tlineL6 1 4
*
* In -----|
* Out -----|
*
Esource      2 0 1 0 1
Tline1      2 0 3 0 Z0=50 TD=10n
Rload1      3 0 50
Eload2      4 0 3 0 1
*
.ends tlineL6
*
.subckt tlineL7 1 4
*
| |

```

```

* In -----| |
* Out -----| |
*
Esource      2 0 1 0 1
Tline1       2 0 3 0 20=50 TD=13n
Rload1       3 0 50
Eload2       4 0 3 0 1
*
.ends tlineL7
*
.subckt tlineL8 1 4
*
* In -----| |
* Out -----| |
*
Esource      2 0 1 0 1
Tline1       2 0 3 0 20=50 TD=10.7n
Rload1       3 0 50
Eload2       4 0 3 0 1
*
.ends tlineL8
*
*****
***** Local Node PLL *****
*
.subckt nodepll 1 2 13 15 9 99
*
* clk -----| | | | |
* clkref -----| | | | |
* output -----| | | | |
* outputbar -----| | | | |
* enable -----| | | | |
*
* When enable is 0V the nodepll is forced to 2.5V
*
Xpfd          1 2 3 4 99 pfd
Xcontroller   3 4 5 6 7 8 9 99 controller
Xchargepump   5 6 7 8 10 99 chargepump
Mpvoltref     10 9 11 99 pmosfet l=1.2u w=10u
Vvoltagegref  11 0 dc 2.46V
Xfilter       10 12 filter
Ebuffer       13 0 12 0 1
Ecomplement1  14 0 12 0 -1
Vcomplement2  15 14 dc 5V
*
.ends nodepll
*
*****
***** Main Circuit *****
*
Vpower        99 0 dc 5V
Vcontrollersig 98 0 pwl(0ns,0V 500ns,0V 501ns,5V)
Vcounterreset 97 0 dc 0V
VDelayBiasPre 96 0 dc 2.5V
VDelayBiasPost 95 0 dc 2.5V
Vnodepllenable 94 0 pwl(0ns,0V 5000ns,0V 5001ns,5V)
*
Xpfd1         18 1 20 21 99 pfd
Xcontroller1  20 21 22 23 24 25 98 99 controller
Xchargepump1  22 23 24 25 26 99 chargepump
Xfilter1      26 27 filter
Xvco1         27 28 99 vco
*
Iload1        28 0 dc 0A
Ebuffer1      29 0 28 0 1
*
Xcounter4bit1 29 97 30 31 32 33 99 counter4bit
Ibit1         30 0 dc 0A
Ibit2         31 0 dc 0A
Ibit3         32 0 dc 0A
Ibit4         33 0 dc 0A

```

```

*
Ebuffer2          1  0 32  0 1
*
XPreDelay_CW_S          1  2 99 fixdelay
Xtline_L1_CW_S          2  3 tlineL1
XDelay_CW_S_N1        3 301 81 201  4 99 capdelayline
Xtline_L2_CW_S          4  5 tlineL2
XDelay_CW_S_N2        5 302 82 202  6 99 capdelayline
Xtline_L3_CW_S          6  7 tlineL3
XDelay_CW_S_N3        7 303 83 203  8 99 capdelayline
Xtline_L4_CW_S          8  9 tlineL4
XDelay_CW_S_N4        9 304 84 204 10 99 capdelayline
Xtline_L5_CW_S         10 11 tlineL5
XDelay_CW_S_N5       11 305 85 205 12 99 capdelayline
Xtline_L6_CW_S         12 13 tlineL6
XDelay_CW_S_N6       13 306 86 206 14 99 capdelayline
Xtline_L7_CW_S         14 15 tlineL7
XDelay_CW_S_N7       15 307 87 207 16 99 capdelayline
Xtline_L8_CW_S         16 17 tlineL8
XPostDelay_CW_S       17 18 99 fixdelay
*
Iview_CW_S_N1          81  0 dc 0A
Iview_CW_S_N2          82  0 dc 0A
Iview_CW_S_N3          83  0 dc 0A
Iview_CW_S_N4          84  0 dc 0A
Iview_CW_S_N5          85  0 dc 0A
Iview_CW_S_N6          86  0 dc 0A
Iview_CW_S_N7          87  0 dc 0A
*
*
XPreDelay_CW_F          29 42 99 fixdelay
Xtline_L1_CW_F         42 43 tlineL1
XDelay_CW_F_N1       43 301 71 201 44 99 capdelayline
Xtline_L2_CW_F         44 45 tlineL2
XDelay_CW_F_N2       45 302 72 202 46 99 capdelayline
Xtline_L3_CW_F         46 47 tlineL3
XDelay_CW_F_N3       47 303 73 203 48 99 capdelayline
Xtline_L4_CW_F         48 49 tlineL4
XDelay_CW_F_N4       49 304 74 204 50 99 capdelayline
Xtline_L5_CW_F         50 51 tlineL5
XDelay_CW_F_N5       51 305 75 205 52 99 capdelayline
Xtline_L6_CW_F         52 53 tlineL6
XDelay_CW_F_N6       53 306 76 206 54 99 capdelayline
Xtline_L7_CW_F         54 55 tlineL7
XDelay_CW_F_N7       55 307 77 207 56 99 capdelayline
Xtline_L8_CW_F         56 57 tlineL8
XPostDelay_CW_F       57 58 99 fixdelay
*
IEnd_CW_F             58  0 dc 0A
*
XPreDelay_CCW_F        117 118 99 fixdelay
Xtline_L1_CCW_F       116 117 tlineL1
XDelay_CCW_F_N1      115 201 121 301 116 99 capdelayline
Xtline_L2_CCW_F       114 115 tlineL2
XDelay_CCW_F_N2      113 202 122 302 114 99 capdelayline
Xtline_L3_CCW_F       112 113 tlineL3
XDelay_CCW_F_N3      111 203 123 303 112 99 capdelayline
Xtline_L4_CCW_F       110 111 tlineL4
XDelay_CCW_F_N4      109 204 124 304 110 99 capdelayline
Xtline_L5_CCW_F       108 109 tlineL5
XDelay_CCW_F_N5      107 205 125 305 108 99 capdelayline
Xtline_L6_CCW_F       106 107 tlineL6
XDelay_CCW_F_N6      105 206 126 306 106 99 capdelayline
Xtline_L7_CCW_F       104 105 tlineL7
XDelay_CCW_F_N7      103 207 127 307 104 99 capdelayline
Xtline_L8_CCW_F       102 103 tlineL8
XPostDelay_CCW_F     29 102 99 fixdelay
*
IEnd_CCW_F           118  0 dc 0A
*

```

```

Xnodepll_N1      71 121 201 301  94  99 nodepll
Xnodepll_N2      72 122 202 302  94  99 nodepll
Xnodepll_N3      73 123 203 303  94  99 nodepll
Xnodepll_N4      74 124 204 304  94  99 nodepll
Xnodepll_N5      75 125 205 305  94  99 nodepll
Xnodepll_N6      76 126 206 306  94  99 nodepll
Xnodepll_N7      77 127 207 307  94  99 nodepll
*
*****
***** Output *****
*
.option post probe
.probe v(1) v(18) v(27) v(28) v(58) v(71) v(72) v(73) v(74)
+ v(75) v(76) v(77) v(81) v(82) v(83) v(84) v(85) v(86) v(87)
+ v(121) v(122) v(123) v(124) v(125) v(126) v(127) v(201) v(301)
*
.ic v(26) = 2.5V
.ic v(27) = 2.5V
.ic v(28) = 0V
*
.tran 0.1ns 10000ns
.end

```

Chapter 7: Conclusion

7.1) Summary

Although gigabit per second electronic backplanes have been demonstrated and some are commercially available, the design of electronic backplanes requires an enormous amount of care to maintain trace line impedance and proper shielding. Similarly, free-space optical designs can require an almost equivalent amount of effort to design the appropriate lens configurations and methods of alignment. Therefore, it may seem that a free-space optical design is an inappropriate alternative especially since the techniques for electrical signal integrity have already been well established. However, free-space optics offers one very important advantage over electronic backplanes; this is its scalability in data rate and parallelism. Not only does it become more difficult to design large parallel electrical systems at ever increasingly higher data rates, but the amount of power required and the cost of these systems will at some point become prohibitive. A point will be reached where the achievable bandwidth of optical interconnects will out-weigh both the design time and cost.

This thesis has described many of the system characteristics for a free-space optical backplane. The optical backplane was offered as a solution to the very short distances of board-to-board and chip-to-chip interconnects. Using a free-space optical design and the 2-dimensional surface area of microelectronic chips, a massive number of parallel optical connections could be provided between printed circuit boards. The data rates at which these optical interconnects could function can easily attain greater than 1-Gbps. This is in part due to the minute optoelectronic devices used and their low loading effects on their respective driving circuitry.

Fundamentally, the optical backplane discussed in this thesis is a collection of point-to-point parallel optical relays with optoelectronic transceivers that are able to convert between optical signals and electrical signals. However, without an interconnect strategy, the design of an optical relay by itself cannot address all the issues concerning a backplane system. Therefore, in chapter 2, a simplified version of the Hyperplane

architecture was introduced. The Hyperplane architecture was used as the underlying interconnection strategy for the optical backplane. Using this architecture, numerous design issues could be addressed, such as the method of packaging, the interaction between the electronics and the optics, the pitch of the printed circuit boards, and most importantly the path each board would use to communicate. The Hyperplane architecture was essentially a generalized, reconfigurable, set of topologies based on a fully connected crossbar interconnect. In this thesis, a particular embedding of the Hyperplane was used. Each board would be responsible for transmitting one channel of data while at the same time it would selectively retrieve data from other optical channels in the backplane. The unit-cell of the Hyperplane architecture was the smart pixel. The smart pixel was capable of three transceiver states; they were called the transparent state, the inject state, and the extract state. Smart pixels were arranged in a 2-dimensional pattern over the surface area of a microelectronic chip such that M-channels of N-bits were formed; this was called a smart pixel array (SPA). Each SPA in the backplane would have control of its own optical channel, which ran through the entire optical backplane. A SPA could determine if data in the optical backplane was destined for it by detecting an address header and comparing it with a stored permanent address.

The physical implementation of the optical design was discussed in chapter 3. Two implementations were discussed, these were known as the Phase-II and Phase-III optical demonstrator systems. Although both systems were based on an uni-directional closed-ring optical design to link 4 PCBs, many features between both designs were different. The Phase-II design was based on a hybrid optical interconnect that combined micro-optics with larger relay lenses. The Phase-III design was a more modular design that integrated sub-assemblies. The Phase-III optical design also used a clustered window per lens approach that allowed the interconnection density to remain high while maintaining an infinite-conjugate optical design style, which allowed easier alignment.

The optical design and the placement of the optoelectronic devices on the microchips were strongly coupled design issues that had to be considered. Certain configurations of optical lens arrangements were not conducive to standard VLSI layout practices. Standard VLSI layout techniques, such as unit-cell repeatability, also did not lead to optimal optical arrangements. Therefore, in chapter 4, design strategies for the

layout of the microchips were developed so that neither the optical design nor the circuit design dominated the layout requirements. An attempt at creating an efficient, repeatable unit-cell structure that also included the optical system requirements was outlined. A structure called a super-cluster was introduced that maintained a streamlined, and efficient circuit design, while allowing the transmit and receive optoelectronic clusters to be separated by a distance of almost 1-mm, resulting in easier optical alignment.

The test results from several optoelectronic microchips were outlined in chapter 5. This chapter showed results from multiple design iterations. Certain insights were obtained through the testing of each microchip, which were not obvious during the design and layout stage. One of the more valuable insights was obtained through failures with the Workshop-Chip, which indicated that internal circuitry could be adversely affected if the optical receiver circuits were left in a floating-state. Fully optical characterizations of the optoelectronic devices indicated that a process error in the flip-chip integration of the optoelectronic devices with the silicon CMOS microchips had occurred. This required structural and process changes to improve the device properties. Information obtained through the design and construction of the testing platforms for the chips was also very helpful when planning the testing procedures of future chip sets. Other data showing the performance of the Phase-III chips indicated that the speed and functionality of the designs were satisfactory and that the chips could be used in the optical backplane demonstrator.

The final part of this thesis dealt with the synchronization aspect of this type of architecture. Whereas most electronic backplanes, regardless of their data rate, regulate the flow of data using bus-controllers and protocols, the architecture proposed in chapter 2 was based on a highly synchronous set of clock signals. This method had the advantage that very little in terms of overhead control is required, since all the operation is governed by regulated periodic signals. However, it is very difficult to control the skew of these periodic signals when the signals must be distributed to a multiple of physically separated points such as PCBs in a backplane. In chapter 6, a set of circuits and a method of balancing the delays among PCBs were presented. This circuit could dynamically compensate for changes in the absolute delay between pairs of PCBs. It would then be possible to regulate a set of clock pulses that would arrive simultaneously at each board

in the system. In this way, an entire system could remain synchronous and there would be no need for complicated protocols that introduce latency.

This thesis has attempted to address some of the design considerations for the introduction of optoelectronic microchips into optical systems by considering both the optical design and the circuit design simultaneously. It has provided useful insight into system-level design considerations through the design and testing of several iterations of the same basic chip architecture. It has provided experimental results for one of the first clustered arrangements of optoelectronic devices on a silicon chip. This thesis has also suggested a novel method of clock synchronization with which to operate a system.

Even though most of the results seem specific to the type of architecture and optical design implemented here, many of the results can extend to other optoelectronic systems because only the most basic and generalized structures were used in the implementation of the chips.

7.2) Future Directions

The next part of the demonstration system will be the integration of the optoelectronic chips with the optical system and the testing of the system as a whole. The quality of the optoelectronic devices will have to be improved, but this is a current topic of study by members of the McGill Photonics Group and Lockheed-Martin Sanders - this problem should be resolved shortly.

During the construction of the Phase-II system, several design changes were devised that would improve the Phase-III system. Unfortunately, the Phase-III system in some ways was so drastically different from the Phase-II system that completely new challenges arose. In a future system, the optical design and the modularity of the system could be improved dramatically. More alignment tolerant relay systems, especially in the chip package assembly, could be realized as well as the placement of more diagnostic structures for the optical system. The flexible printed circuit board that carried the optoelectronic chip would be eliminated and methods of fixing the optoelectronic chip to a removable hard PCB would be considered.

One of the most crucial design changes would be to eliminate the modulator optoelectronic technology and replace it with emitter-based technology such as vertical cavity surface emitting lasers (VCSEL). This would simplify the optical design in some ways, since an external optical power supply would not be required. However, this would once again change the style and layout of the microchip. Issues such as the temperature stabilization and the amount of power required by arrays of VCSELs would have to be addressed.

Finally, the amount of circuitry on the microchips would be increased to include the ability to buffer optical data as well as better methods for re-timing data and interfacing to external circuitry. The microchips built thus far have included only the basic minimum amount of circuitry required to demonstrate the bandwidth of the optical backplane. Future chips would perhaps be built based more closely on an industry standard such as SONET or ATM.

This technology will evolve and diversify in the future, and free-space optical interconnects may start to move out of the application specific areas and into more mainstream interconnect technology as current electronic switching technology can no longer be adapted to greater data rates. However, it will be doubtful that the systems discussed in this thesis will be immediately adopted by industry. Slow and thoughtful steps towards the "free-space alternative" will be made where small subsections of systems will at first benefit from optics and as the confidence in these subsections increases, more of the entire system will evolve towards the fully optical interconnect.