Process Based Cost Modeling of Emerging Optoelectronic Interconnects: Implications for Material Platform Choice

by

Shan Liu

Bachelor of Science, Electrical Engineering, University of Texas at Austin, 2006

Submitted to the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Technology and Policy

at the

Massachusetts Institute of Technology

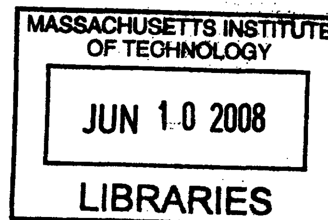
June 2008

©2008 Massachusetts Institute of Technology.
All rights reserved.

Signature of Author..
Technology and Policy Program, Engineering Systems Division
May 15, 2008

Certified by..
Randolph E. Kirchain, Jr.
Assistant Professor of Material Science & Engineering and Engineering Systems
Thesis Supervisor

Accepted by..
Dava J. Newman
Professor of Aeronautics and Astronautics and Engineering Systems
Director, Technology and Policy Program



ARCHIVES

Process Based Cost Modeling of Emerging Optoelectronic Interconnects: Implications for Material Platform Choice

by

Shan Liu

Submitted to the Engineering Systems Division on May 15, 2008
in Partial Fulfillment of the Requirements for
the Degree of Master of Science in Technology and Policy

Abstract

Continuously increasing demand for processing power, storage capacity, and I/O capacity in personal computing, data network, and display interface suggests that optical interconnects may soon supplant copper not only for long distance telecommunication but also for short reach connection needs. In the search for a standard, the current debate in the optoelectronic industry is focused on the technical and economic challenges of the next generation interconnect. Technological advances over the past few years have given new strength to a silicon-technology platform for optoelectronics. The possibility of extending a mature and high-yield Si CMOS manufacturing platform of the electronic industry into the optical domain is an area of intensive interest.

Introducing new photonic materials and processes into the mature electronic industry involves a convergence of knowledge between the optoelectronics and semiconductor IC manufacturers. To address some of the technical, market, and organizational uncertainties with the Si platform, this research explores the economic viability and operational hurdles of manufacturing a 1310 nm, 100G Ethernet LAN transceiver. This analysis is carried out using the process-based cost modeling method. Four transceiver designs ranging from the most discrete to a high level of integration are considered on both InP and Si platforms. On the macro-level, this research also explores possible electronic-photonic convergence across industries through a multi-organization, exploratory roadmapping effort.

Results have shown 1) integration provides a cost advantage within each material platform. This economic competitiveness is due to cost savings associated with the elimination of discrete components and assembly steps; 2) a total cost comparison across material platforms indicates at low volume (less than 1.1 million annual units), the InP material platform is preferred, while at high volume (greater than 3 million annual units) the Si material platform is preferred. Furthermore, this study maps out the production cost at each technology and volume projection, and then compares this cost with price expectation to determine the viability of the transceiver market in the datacom and computing industry. Results indicate that annual production volumes must be in the tens of millions unit range to provide the minimum economies of scale necessary for designs to meet the trigger price. These results highlight that standards and a set of common language are essential to enable converging technology markets.

Thesis Supervisor: Randolph E. Kirchain, Jr.
Assistant Professor of Material Science & Engineering and Engineering Systems

Acknowledgement

I would like to express my deepest gratitude to the people who made my two years at MIT the most memorable and intellectually stimulating years of my life. This thesis work would not have been possible without the support and guidance from my mentors, friends, and family.

First and foremost, I am extremely lucky to have Dr. Randy Kirchain as my thesis supervisor. Randy's high expectation of my work from an early stage pushed me to strive for the best quality research. Randy's openness, intellectual acumen, hard working ethics, and humor all made my time at the Materials Systems Lab (MSL) both intellectually challenging and fun. Randy encouraged me to dig below the surface of many established theories, and ask meaningful as well as impactful thesis questions. In addition, I am particularly grateful for Randy's meticulous review of my thesis draft and providing insightful comments.

Along this research journey, I am very thankful to Dr. Erica Fuchs. She walked with me every step of the way in my first year, and provided generous support during the second year. Erica is an amazing mentor and friend. She provided essential research directions and industry contacts that made this work possible. I learned tremendous research skills from her that are invaluable for my future career. Erica spent countless hours with me both in and out of the office. She showed me the meaning of a great researcher who always asks the hard questions.

I would also like to thank the people at the MIT Microphotonic Center, in particular, Prof. Lionel Kimerling, Mark Beals, and Mindy Baughman. Prof. Kim and Mark shared with me their deep technical knowledge on photonic technology as well as first-hand industry experience. Mindy was always a great help to my research by informing me about previous work done at the MPC. They also gave me opportunities to present this work in four MPC consortium meetings. In addition, I am grateful to all of my industry contacts at the Si Technology Working Group that made this research possible.

Equally importantly, people at MSL, in particular, Dr. Richard Roth, Dr. Frank Field, and Terra Cholfin helped me tremendously. Rich and Frank always asked the key questions at the right time. They could find any weakness of my research and their guidance often saved me hours of pointless work. My fellow students at MSL, Elisa, Trisha, Sup, Marie-Claude, Gabby, Yingxia, Catarina, Susan, Tommy, Anna, Rob, Ashish, Jonathan, and Yaoqi are wonderful friends that supported me through the good and hard times. MSL is a wonderful community that I will surely miss in the future.

I am extremely lucky to have received an education from the Technology and Policy Program (TPP) at MIT. I would like to thank Prof. Dava Newman, Frank Field, and Sydney Miller for believing in me and continuously supporting my education. I would also like to thank my professors, Kenneth Oye, Stuart Madnick, and Jason Davis, for their excellent course materials that I incorporated into my thesis. Furthermore, I am very grateful to be friends with all 40 of my fellow TPPers, for the amazing work they do and contributions they will make in the world.

Finally and most importantly, I thank my parents, Lijun Yuan and Boying Liu. Their love, courage, sacrifice, and wisdom are the source of my strength and inspiration in life. With their unwavering caring and support, nothing seems impossible. My gratitude to them could not be sufficiently expressed in words.

Table of Content

Abstract ... 3
Acknowledgement ... 5
Definitions ... 13
1 Introduction ... 15
1.1 Motivation ... 15
1.2 Problem Statement ... 18
1.3 Research Questions ... 21
1.4 Process-Based Cost Modeling Method ... 21
1.5 Contribution and Chapter Outline ... 22
2 Background ... 23
2.1 Ethernet Standard—A Short History and Recent Development ... 23
2.2 State of Art Designs for 100G Ethernet LAN Transceiver ... 25
2.2.1 Competing Material Platform ... 27
2.2.1.1 III-V Group ... 28
2.2.1.2 Silicon ... 28
2.2.2 Competing Design ... 29
2.2.2.1 System Architecture ... 29
2.2.2.2 Integration Scheme ... 31
2.3 Current State of the Optoelectronic Industry and the Road Ahead ... 34
2.4 The Importance of Standards and Coordination ... 36
2.4.1 Introduction to Roadmapping ... 36
2.4.2 Lessons from the Semiconductor Industry ... 37
3 Method ... 41
3.1 Existing Cost Modeling Methods ... 41
3.2 Process-Based Cost Modeling ... 44
3.2.1 Conceptual Framework ... 44
3.2.2 Model Architecture and Capability ... 45
3.3 Building Virtual Manufacturing Facilities ... 54
3.3.1 III-V Fab ... 54
3.3.2 Si Fab ... 55
3.3.3 Backend ... 55

3.4 Data Collection 56
3.5 Roadmapping Method: Cross Market Interviews....................56
3.5.1 General Roadmapping Methodologies 57
3.5.2 Cross Market Questionnaire 58
4 Results 59
4.1 Layout of Four Designs 59
4.1.1 Design 1. TO-CAN 59
4.1.2 Design 2. DML Array....................60
4.1.3 Design 3. Hybrid 61
4.1.4 Design 4. Si Two Chip 61
4.1.4.1 One Chip Solution 62
4.2 Cost Analysis on Integration 63
4.2.1 InP Designs 63
4.2.1.1 Frontend (Lasers) Cost Comparison 63
4.2.1.1.1 Economy of Scale....................64
4.2.1.1.2 Top Cost Drivers by Process Types and Cost Element 65
4.2.1.1.3 Sensitivity of Cost Estimate to Production and Process Yield 66
4.2.1.2 Backend comparison....................69
4.2.1.2.1 Economy of Scale....................69
4.2.1.2.2 Top Cost Drivers by Process Type and Cost Element....................70
4.2.1.3 Conclusion 71
4.2.2 Si Designs....................73
4.2.2.1 Frontend (Photonic Chips) Comparison 73
4.2.2.1.1 Economy of Scale....................74
4.2.2.1.2 Top Cost Drivers by Process Types and Cost Element 74
4.2.2.1.3 Sensitivity of Cost Estimate to Production and Process Yield 75
4.2.2.2 Backend Comparison....................77
4.2.2.2.1 Economy of Scale....................78
4.2.2.2.2 Top Cost Drivers by Process Types and Cost Element 78
4.2.2.3 Conclusion 79
4.2.2.4 Cost Estimation on Si One Chip Design....................80
4.3 Cost Comparison of Material Platforms 81
4.4 Analysis on Three Factor Comparison 84
5 Conclusion and Future Work 88
5.1 Research Conclusions 88

5.2 Thesis Contribution and Future Research 91
Reference 93
Appendix A. Background Information on IEEE HSSG 97
Appendix B. Optical Components – Computing Market Segment Questionnaires 99
Appendix C. Benchmarking with External Models........ 101
Appendix D. Analysis on Wafer Size Transition 102

List of Figures

Figure 1. Global optoelectronics 10 years forecast for component segments 15
Figure 2. Trend in the information-carrying capacity of a single line 16
Figure 3. Market transition: electrical-photonic convergence 17
Figure 4. S-curve for the 100GE transceiver market 20
Figure 5. Generic Ethernet Physical Layer Reference Model 25
Figure 6. Possible levels of integration for 100GE transceivers 31
Figure 7. a) A 10x10G DWDM solution, b) Photonic Integrated Circuit, transmit and receive chips, c) Discretely packaged 100GE transmitter and receiver 33
Figure 8. Photonic Integrated Circuit, a CWDM solution—100GE transmitter in a single package 33
Figure 9. Comparison of electronic and photonic roadmapping on S-curves 36
Figure 10. Conceptual framework of PBCM 45
Figure 11. FOP-PBCM Structure 45
Figure 12. Category of time breakdown based on line utilization for a 24 hour day 50
Figure 13. Stages of technology roadmapping 57
Figure 14. TO-CAN 60
Figure 15. DML Array 60
Figure 16. Hybrid 61
Figure 17. Si Two Chip 62
Figure 18. 10 by 10G discrete laser vs. 100G DML array cost comparison (frontend) 64
Figure 19. Discrete lasers frontend top seven processes driving cost at 1 million APV 65
Figure 20. DML array frontend top seven processes driving costs at 1 million APV 66
Figure 21. 10 by 10G discrete laser vs. 100G DML array cost comparison, frontend 67
Figure 22. Discrete 10G laser cost vs. total yield, frontend 67
Figure 23. TO-CAN frontend cost elasticity 68
Figure 24. DML array frontend cost elasticity 68
Figure 25. 10 by 10G discrete lasers vs. 100G DML array cost comparison, backend 70
Figure 26. Discrete 100G backend top five processes driving cost at 1 million APV 71
Figure 27. 100G DML array backend top five processes driving costs at 1 million APV 71
Figure 28. 100G TO-CAN vs. DML array total cost comparison 72
Figure 29. Hybrid vs. Si Two Chip, frontend photonic chip cost comparison 74
Figure 30. Hybrid frontend top seven processes driving cost at 10 million APV 75
Figure 31. Si Two Chip frontend top seven processes driving cost at 10 million APV 75
Figure 32. a) Hybrid total cost, sensitivity of APV to total yield, and b) Si Two Chip total cost, sensitivity of APV to total yield 76
Figure 33. Hybrid frontend cost elasticity 77
Figure 34. Si Two Chip frontend cost elasticity 77
Figure 35. Hybrid vs. Si Two Chip cost comparison, backend 78
Figure 36. Hybrid backend top five processes driving cost at 10 million APV 79
Figure 37. Si Two Chip backend top five processes driving cost at 10 million APV 79
Figure 38. Hybrid vs. Si Two Chip total cost comparison 80
Figure 39. a) Total cost comparison of four designs, b) delta cost using Si Two Chip as the baseline at high volume 82
Figure 40. Cost comparison of four designs, backend 83

Figure 41. Three Factor Analysis 84
Figure 42. Trigger price comparison delta graph, datacom 86
Figure 43. Trigger price comparison with Si Two Chip (delta), computing 87
Figure 44. IEEE Structure 97
Figure 45. IEEE HSSG 100GE Timeline 97
Figure 46. 4x25G Transceiver Architecture for Metro Applications includes SOA(s) 98
Figure 47. Wafer size lifecycle 102
Figure 48. Hybrid frontend cost comparison, 200 mm vs. 300 mm 104
Figure 49. Si Two Chip frontend cost comparison, 200 mm vs. 300 mm 104
Figure 50. a) Si Two Chip 200mm wafer die cost, total cost: $17.5, b) Si Two Chip 300mm wafer die cost, total cost: $14.2 at 18.55M APV 105

List of Tables

Table 1. Material alternatives for transceiver components 27
Table 2. Technology-Architecture alternatives 30
Table 3. Technology-Architecture match proposed by APIC 30
Table 4. Discrete 100GE optical solution one 32
Table 5. Discrete 100G optical solution two 32
Table 6. OE components industry business model 35
Table 7. A Comparison between the Optoelectronic and the Semiconductor industry 40
Table 8. High level operational parameters on general input sheet 46
Table 9. Process Flow Input Table 47
Table 10. Process Recipe Input 48
Table 11. Final result table of FOP-PBCM in Cost Summary sheet 53
Table 12. Process Module and representative Process Recipe for III-V Fab 54
Table 13. Process Module and Process Recipe for Si Fab 55
Table 14. Backend Process Module 55
Table 15. Assumptions in the frontend fabrications of Design 1 and 2 63
Table 16. Backend yield assumptions in Design 1 and 2 69
Table 17. Frontend model assumptions for Design 3 and 4 73
Table 18. Cross market survey results, datacom 85
Table 19. Server volume projection 85
Table 20. Transceiver volume, trigger price, and PBCM cost results across four designs 85
Table 21. PBCM cost results across four designs minus trigger price 85
Table 22. Cross market survey results, computing (chip-to-chip). 86
Table 23. Three model results comparison for Design 4 in an 8 inch Si fab 101
Table 24. Frontend model assumptions for 200mm and 300mm wafer size 103

Definitions

AWG arrayed waveguide
APV annual production volume
CMOS complementary metal oxide semiconductor
COO Cost of Ownership
CTR Communication Technology Roadmap
CWDM coarse wavelength-division multiplexing
DBR distributed bragg reflector (laser)
DFB distributed feedback (laser)
DML directly modulated laser
DWDM dense wavelength-division multiplexing
EAM electro-absorptive modulator
EML electro-absorption modulated laser
FAB fabrication facilities
G gigabits
GE gigabits Ethernet
IC integrated circuits
LAN local area network
MIT Massachusetts Institute of Technology
MPC Microphotonic Center/Consortium
MZI Mach-Zehnder interferometer
OE optoelectronic
PBCM process-based cost modeling
PC photonic chips or personal computing
PD photo-diode
PIC photonic integrated circuit
PIN positive intrinsic negative (diode)
PLC planar lightwave circuit
ROSA receiver optical subassembly
SOI silicon-on-insulator
TIA transimpedance amplifier
TOSA transmitter optical subassembly
TWG Technology Working Group
TWPD traveling-wave photo-detector
VCSEL vertical-cavity surface-emitting laser
VLSI very large scale integration
WDM wavelength-division multiplexing
XENPAK multi-source agreement defining a 10GE transceiver form-factor

1 Introduction

1.1 Motivation

The past twenty years have transformed our society from a natural resource based economy to a knowledge based one. The life support of this knowledge economy is an information infrastructure, where network and personal computing enable the seamless transfer of intangible information, in bits and byte, to millions of interconnected citizens. Furthermore, the new millennium has observed the dawn of an advanced information age, which is characterized by the explosion of user generated content through the emergence of Web 2.0, ubiquitous access, and sensor networks, as well as improvement in virtual reality, display quality, and real-time video processing. These new media applications drive an unprecedented demand for bandwidth, information capacity, and computing power. This demand is manifested in the various forecasts for global optoelectronics components—one of the core technologies of communication. These forecasts project steady growth with consumer entertainment a strong driver (e.g., Figure 1). Have we only seen the tip of the iceberg in the biggest information revolution of our time?

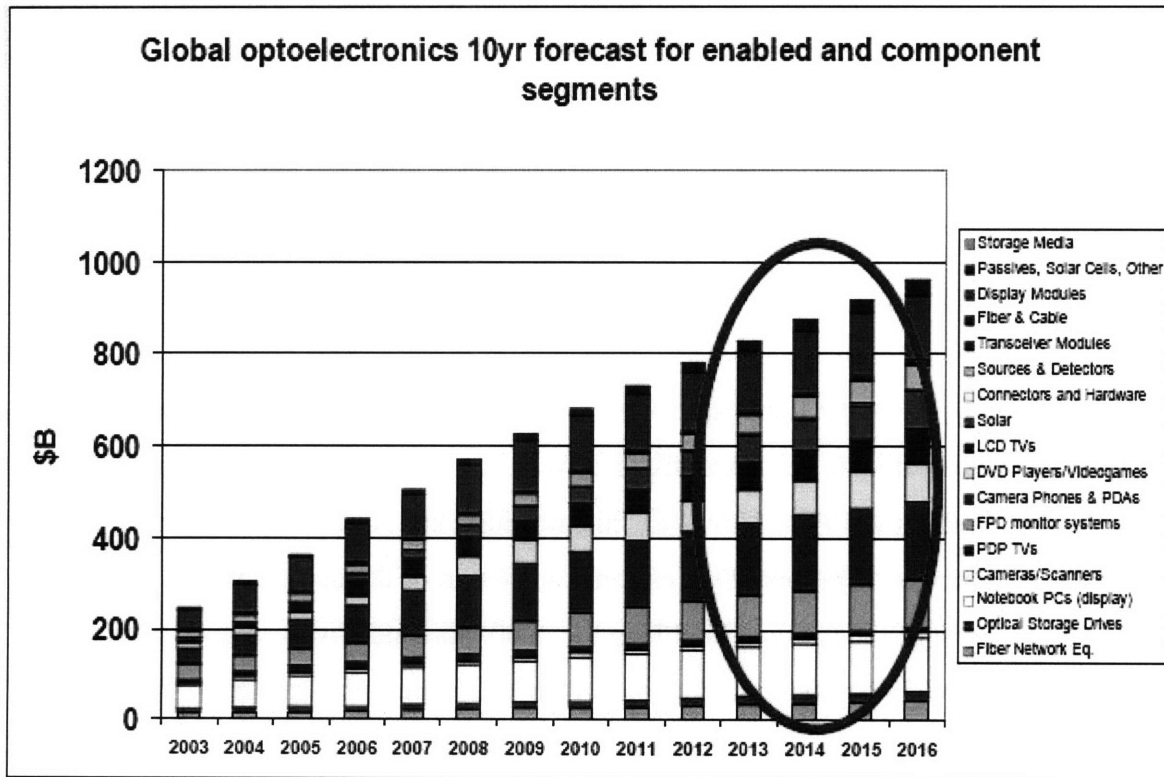


Figure 1. Global optoelectronics 10 years forecast for component segments (Lebby 2006)

Given this projected future demand for interconnection capacity, many are concerned that copper interconnects in the datacom and computing industry will likely reach technical limitations in the near future. At high bandwidth, copper interconnects are affected by loss and crosstalk. Some experts believe that copper's limit is in the 10 to 40 GHz range for PC boards at less than 1 meter distance (Bautista, Morse et al. 2005). There is an increasing need for higher bandwidth that copper may not be able to meet. Optical interconnects may be the ideal substitute. Relative information capacity per line of optical fiber is higher than copper coaxial cable. Figure 2 shows the increase in information capacity per transmission line through time.

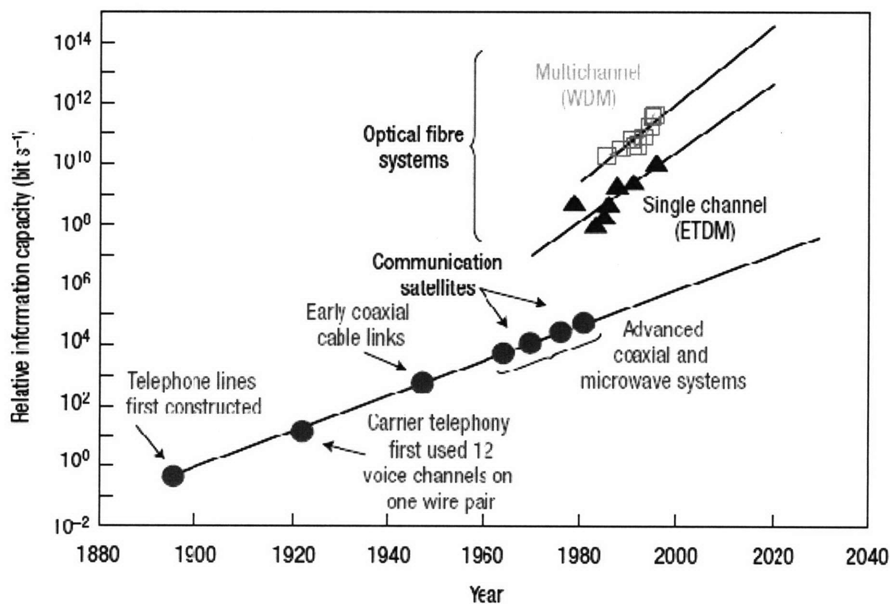


Figure 2. Trend in the information-carrying capacity of a single line (Kirchain and Kimerling 2007)

Transitioning from copper to optical solutions has faced many barriers. The replacement of copper wire by optical fiber is one disruptive innovation the telecommunication industry has successfully adopted in the 1990s. However, this transition has gone from a period of "Irrational Exuberance" to a time of "Irrational Depression" on the part of the optoelectronic industry (Cole 2007). After the burst of the dot com bubble in the 2000, burdened by overcapacity and overspending, the optoelectronic industry began to hold a more cautious attitude that weighs cost minimization over radical innovation (Fuchs, Bruce et al. 2006).

In short-distance applications for datacom, the industry is observing a gradual transition from copper to optical interconnects. Figure 3 maps out a possible scenario of such transition occurring in the next 5 years.

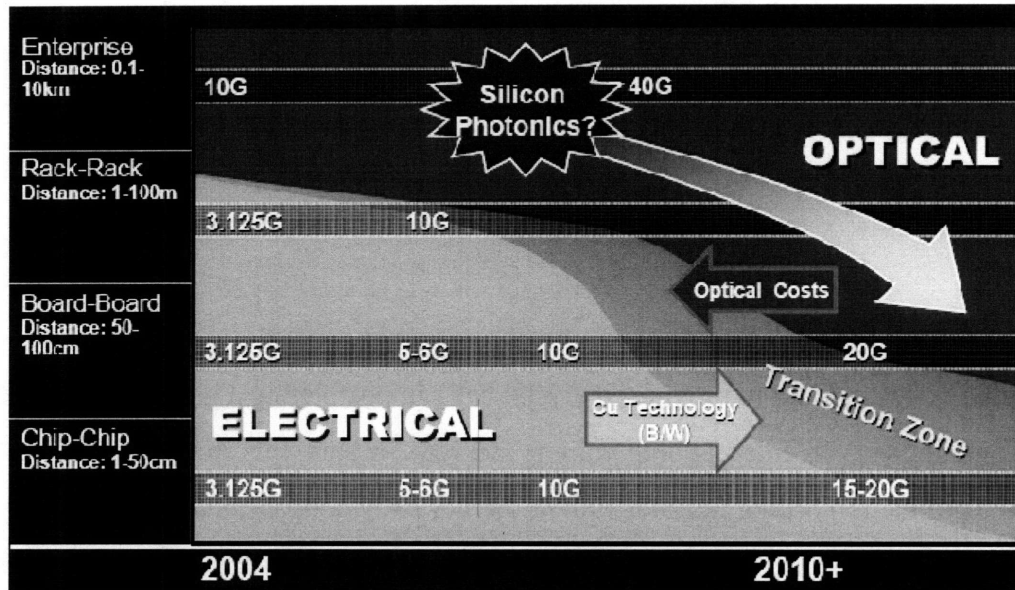


Figure 3. Market transition: electrical-photonic convergence
Courtesy of Intel (Bautista, Morse et al. 2005)

The Computing industry has also taken an initial interest in optical interconnects due to concerns over the continuation of Moore's Law. "The semiconductor industry is no longer scaling at its historical rate as industry seemingly approaches the end of improvement rates aligned with the predictions of Moore's Law" (Bruce and Fine 2007). A communication bottleneck is threatening VLSI electronics at data rates problematic for copper interconnects. For example, the eight-core processor in Sony's Playstation 3 game console has a computation power of 256 billion floating point operations per second, and it communicates with the peripheral graphics processor and memory at data rates of 25G or higher (Jalali and Fathpour 2006) – already within the range of concern for copper interconnection technology. Electrical-photonic integration at the chip level may be able to alleviate this information bottleneck by creating a new platform—the planar waveguide, in the micro- and nano- scale world (Kirchain and Kimerling 2007).

All the above technological and market developments bring forth opportunities for the optoelectronic industry to enter new markets in the next 5 to 10 years. As the numbers of potential applications for photonics grow, the optoelectronic industry faces many new

challenges. These challenges include the lack of or limited common standards and a common manufacturing platform, component down-pricing pressure, lack of a foundry model, absence of large-scale industry coordination, and proliferation and diversification of products (Bruce and Fine 2007). A framework for thinking about the future of this industry is needed to identify potential disruptive innovations, explore alternative technological paths, and establish industry coordination. This research is motivated by the danger and opportunity faced by the optoelectronic industry, and the potential catalyzing societal change integrated photonic circuits may bring.

1.2 Problem Statement

This study does not aim to provide "a crystal ball" in postulating a general strategy for the optoelectronic industry in entering high volume markets. Instead, this study uses a specific case study on the production economics of a 100G Ethernet LAN optical transceiver to study cost competitiveness of various photonic solutions. The 100GE transceiver provides an insightful case study because it is a hotly debated emerging technology that engenders a variety of designs across various material platforms (III-V, Si, and hybrid) and levels of integration. Furthermore, the transceiver is used as an entry point for the discussion of an industry roadmap since it is a common component across industries.[1] The audiences for this research include the traditional optoelectronic component manufacturers, as well as established players from the computing industry. In the short term, it may be advantageous for the computing industry to take a foothold in the 100G Ethernet space as a test bed for photonics on the Si platform.

The school of innovation literatures defines a framework of sustaining verses disruptive technology. Sustaining technologies can be either incremental or radical. Both types improve product performance by reinforcing an established technology trajectory. Disruptive technologies often result in worse product performance in the near term, but in the long term, they are often cheaper, simpler, and smaller than the incumbent products. Disruptive technologies are typically commercialized by new firms rather than established market leaders, and are often targeted for an emerging market, but later can move up-market by displacing the dominate sustaining technologies. (Christensen 1997)

Optical transceivers are currently commercialized on the III-V material platform. III-V is the

[1] Discussion from the MIT CTR Cross-market TWG

sustaining technology that may remain to be the material of choice for photonic integration due to its superior lasing property (Clayton and Dudley 2005). However, technological advances over the past few years have given new strength to a silicon-technology platform for optoelectronics. The possibility of extending a mature and high-yield Si CMOS manufacturing platform into the optical domain is an area of intensive interest due to its low cost potential from optical-electronic integration and high volume manufacturing capabilities. Si photonics may become a source of disruptive innovation to the incumbent III-V transceivers.

Technology S-curves is one useful framework to study technology evolution. It shows a relationship between a product's performance improvement and the engineering effort required to realize that performance. The shape of the curve suggests that as the technology matures, the rate of progress slows down and eventually approaches a natural limit (Christensen 1997). Given this trend, in technology development, there exists a diminishing return of performance improvement gained to engineering efforts spent. Figure 4 shows a possible S-curve scenario for the III-V and Si photonic transceiver markets. Performance on the y-axis can be measured by production yield or transceiver unit revenue (inverse of cost). III-V transceivers as the sustaining technologies can be plotted as a series of intersecting S-curves. Discontinuity occurs if integrated III-V transceivers can surpass the more incremental technology—discrete III-V transceiver in performance. In addition, Si photonics based transceivers as the disruptive innovation can be conceptualized on a separate S-curve plot because they are initially measured by a different set of performance requirements in a separate market. The key questions are: 1) whether the disruptive technology, Si photonic in this case, has the potential to displace the more established III-V technology and invade multiple technically converging markets, and 2) what roles standards should play to assist such technology adoptions.

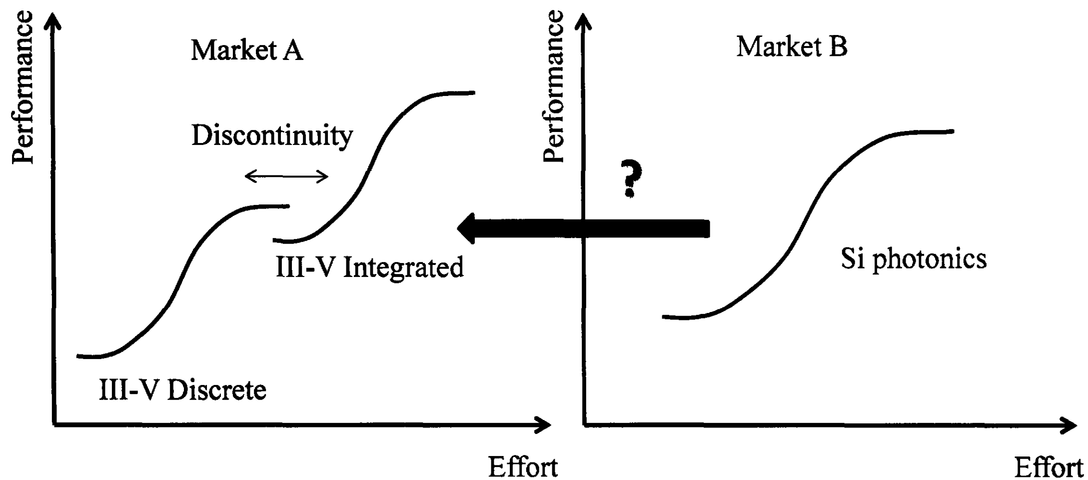


Figure 4. S-curve for the 100GE transceiver market

Difficulties in realizing Si photonics' potential are results of three major uncertainties with a Si platform[2]:

1) Achieving a technically viable device
2) Achieving an economically competitive device
3) Achieving a superior manufacturing capability

To begin to answer these questions, this research investigates the manufacturing cost of emerging InP and Si designs for a 1310 nm, single mode, 100G Ethernet LAN transceiver for less than 300 meter distance applications. There are a variety of 100GE solutions being considered in industry. To better understand the impact of integration on production costs, four designs are investigated: (1) a discretely packaged InP transmitter with 10 TO-CAN & discretely packaged receiver, (2) a discretely packaged InP transmitter with 10 by 10G directly modulated laser (DML) array & discretely packaged receiver, (3) a hybrid transceiver with InP laser and detector array, and (4) a monolithically integrated optical die & InP DFB laser array in a single package.

. Integrated photonics are at an early stage of development. As mentioned in the motivation section, the challenges faced by the optoelectronic industry create many roadblocks for integration and high volume manufacturing. A possible solution is learning from the semiconductor industry. The optoelectronic industry may benefit from a common vision and cohesive plan using the semiconductor industry's International Technology Roadmap for

[2] Parts of these ideas are from a conversation with Dr. Erica Fuchs, CTR Fellow, Spring 07.

Semiconductors (ITRS) as a role model (Bruce and Fine 2007). This research explores possible electronic-photonic convergence across industries by participating in the MIT Communication Technology Roadmapping program. This is a multi-organization, exploratory roadmapping effort to identify potential standard and organizational changes.

1.3 Research Questions

This research aims to answer the following questions:

1) Technology: What is the most cost competitive architecture and material solution in manufacturing 100G Ethernet LAN transceivers?

2) Market: What is the structure of the three factors: cost of achieving the technology, volume expectation, and required substitution price that will characterize an advantageous state for Si photonics to enter the datacom and computing markets?

3) Organization: What roles could industry standards and coordination (roadmapping) play as enablers of emerging optoelectronic interconnect technology across markets?

1.4 Process-Based Cost Modeling Method

Previous studies on the cost-feasibility of emerging interconnect technologies are either limited in scope—focusing on cost advantages of further integration for a particular device in a specified material system (Fuchs, Bruce et al. 2006), or limited in the computational tools that fail to integrate cost analysis into product and process development. This research uses the process-based cost modeling method (PBCM). PBCM is best suited for mapping the intricate details and consequences of design and manufacturing processes into a quantifiable cost metric. Alternations in a product's architecture and material platform have real consequences leading to different operating conditions and yield. Therefore, a prospective rather than a retrospective cost modeling method is better used for studying the economic viability of emerging technology (Field, Kirchain et al. 2007). The author built on a previously available PBCM based model from the MIT Material System Lab by adding modeling capabilities for semiconductor processing. The detail of the PBCM will be explained in the Chapter 3 of this thesis.

1.5 Contribution and Chapter Outline

Analysis from this research can help to provide an understanding of the technical, economic, and organizational challenges for disruptive emerging optoelectronic technology to gain commercial acceptance, and the possible spillover effects of this new knowledge to the datacom and computing industry. These results can provide insights and directions for firms' technology strategy and development trajectory toward a higher degree of device integration on both III-V and Si platform, as well as a motivation to reach common standards across markets.

Chapter 2 provides some background in the development history of the Ethernet leading to 100GE specification, reviews the state of art in competing transceiver designs, current state of the optoelectronic industry, and its differences from the more matured semiconductor industry. Chapter 3 first conducts a literature search in existing cost modeling methods, and then presents the methodology of this research in terms of building virtual fabrication facilities, data collection, and cross market survey in the MIT CTR's roadmapping effort. Chapter 4 is the climax of this thesis, containing 100GE transceiver cost modeling analysis and results of the MIT CTR cross market survey. Chapter 5 concludes with comments on technology, market, and organizational barriers for Si photonic technology to gain commercial success. This final chapter also revisits thesis contributions and delineates a prospect for future research.

2 Background

This chapter provides a brief history of Ethernet LAN technology and standard development leading to the most recent 100GE preliminary specification. Next, a survey of emerging 100GE transceiver designs is presented. In order to capture value from economies of scale in manufacturing, firms may need to establish several standard designs to reach the desired "high volume—low cost" target. However, competing designs in III-V and Si material platforms at various integration levels present a challenge for standardization. Next, this section presents the current state of the optoelectronic industry. The semiconductor industry is used as a success story to highlight the importance of standard and coordination, and potentials for knowledge transfers between the two converging industries.

2.1 Ethernet Standard—A Short History and Recent Development

In the 1970s, Xerox Corporation developed Ethernet as a coaxial cable network. The first experimental system operated at a data rate of 3 Mbps using a carrier sense multiple access collision detect (CSMA/CD) protocol. Ethernet became a commercial-quality system in 1980 through a joint development effort of Digital Equipment Corporation, Intel, and Xerox. Their 10-Mbps Ethernet Version 1.0 specification became the forerunner for the IEEE 802.3 standard, which was approved in 1983 and officially published in 1985. Since 1985, all subsequent Ethernet equipment was built according to the IEEE 802.3. This standard is periodically updated to support newer network media and higher data rate, as well as new network access control features. (Ford and Cisco Systems Inc. 1997) Currently, 10G Ethernet is the fastest existing IEEE Ethernet standard.

In recent years, demand for bandwidth is increasing in many communication market segments. Advanced media applications, such as interactive TV and on-line gaming, enable higher user interaction and system interoperability, and therefore drive large demand for upstream and downstream channel bandwidth. According to Comcast, the shift from Broadcast to Unicast (personalized and interactive) services[3] would drive core network capacity. The current forecasts project that by year 2011 Narrowcast and Unicast services will be greater than 20% of all service offered (Saxena 2007). With the emergence of Web 2.0 and new interactive

[3] Broadcast services include analog video, digital broadcast, and digital simulcast. Narrowcast and Unicast services include DOCSIS, and Video-On-Demand.

services on the horizon, enterprise, service and content providers have begun to search for the next generation of gigabits Ethernet technology.

The current generation of high speed Ethernet technology, 10GE, began to see wide adoption in 2007, but is already considered obsolete to satisfy future bandwidth-intensive applications. A new IEEE Higher Speed Study Group (HSSG) formed in the summer of 2006 aims to publish a 100GE specification by 2010[4] (Wirbel 2006). The industry is focusing on what comes after aggregating 10GE pipes. Parallel nx10G links have limitations on balancing load distribution. Questions have been raised on the number of parallel 10G links needed to match usable bandwidth on one 100G link (Saxena 2007).

Presently, the optoelectronic industry is seeking a standardized and cost effective solution to facilitate 100GE development and adoption. Commercializing 100GE is expected to be much more difficult than the 1G to 10G transition. The IEEE 802.3 HSSG interim meetings from 2006 to mid-2007 showed a lack of consensus on a low cost, technologically feasible transceiver design. In July 2007, the group submitted a Project Authorization Request to the 802 Standards Executive Committee for the approval of a new IEEE 802.3ba standard, which includes both 40G and 100G data rates operating over optical fiber and copper cable. The physical layer specifications supporting 100GE will operate over single-mode fiber for distances up to 40 km. Technology selection begins in early 2008 and a last round of new proposals will be accepted until early summer[5] (HSSG). This research positions itself right into the heart of the current technology debate.

[4] Figure 44 in Appendix shows the structure of IEEE standard group

[5] Figure 45 in Appendix shows the timeline for 100GE standard formulation

2.2 State of Art Designs for 100G Ethernet LAN Transceiver

Before delving into details of transceiver designs, this section first presents a high level overview of the IEEE 802.3 architecture. The Open Systems Interconnection Basic Reference Model (OSI Model) is used as a blueprint for communication and computing network protocol. The IEEE 802.3 standard group defines the media access control (MAC) sub-layer of the data link layer, and the physical layer of the OSI Model, as shown in Figure 5. Wired Ethernet devices only implement the bottom two layers on the OSI Model stack, in the forms of network interface cards that can be plugged into a host device's motherboard. Physical layer attributes are transmission rate, transmission method, and the media type/signal encoding. This study primarily concerns with the physical medium attachment (PMA) sub-layer, which contains the transceiver; as well as the media-dependent physical coding sub-layer (PCS), which includes multiplexing and demultiplexing of data streams. (Ford and Cisco Systems Inc. 1997)

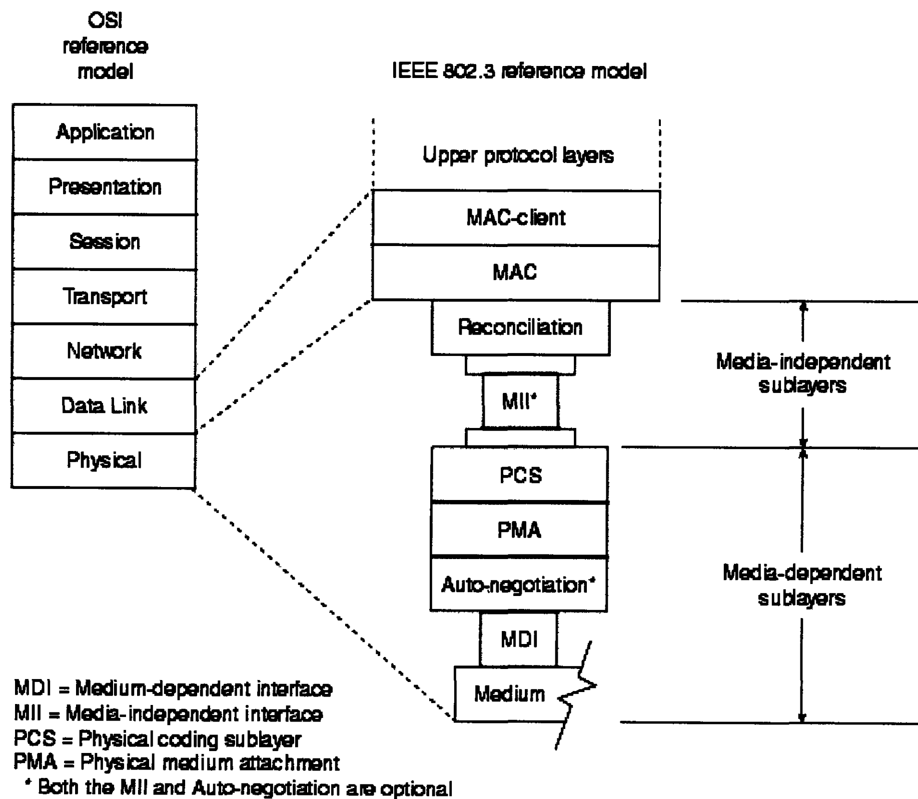


Figure 5. Generic Ethernet Physical Layer Reference Model
(Ford and Cisco Systems Inc. 1997)

The author conducted interviews with several large optical components firms during Fall 2006. Results from these early interviews confirmed that a common direction for 100GE transceivers has not yet emerged in the 2006 to 2007 timeframe. Companies showed diverse approaches to material platform and design architecture, ranging from:

–"We currently still are considering all options—10x10G, 4x25G, 2x50G, and 1x100G"–"We have decided some sort of photonic integrated circuit (using some amount of monolithic integration) is definitely the way to go."–"In the short-term, we believe a hybrid solution combining InP and Si will need to be used in our product. Whether in the long-term we will turn to all-silicon is still up in the air."–"The answer is clear. InP has already demonstrated low cost manufacturing of 100G parts."
(Liu and Fuchs 2007)

As seen from recent HSSG discussions in 2008, industry consensus has tilted toward a 4 x 25G WDM architecture for a 1312 nm wavelength center transmitter (Appendix A, Figure 46), as well as continuous discussion on a low cost, 1310 nm un-cooled CWDM DML solution for the near term. The next two sub-sections present a survey of emerging transceiver designs in various competing material platforms and design architectures.

2.2.1 Competing Material Platform

Two key material alternatives for optoelectronic device fabrication are Si and III-V materials (Al, Ga, As, In, and P). Available material choices for each transceiver component are shown in Table 1. The III-V material system is the traditional platform to build optical devices due to its efficiency in generating light and capability in a full range of photonic functions. As a result, in contrast to Si platform's lack of lasing capability, the III-V platform is a natural choice in which to pursue photonic integration (combining active and passive optical elements). However, electronic-photonic integration will become increasingly important at data rates well beyond 10G, Si then may become the material choice at high degrees of integration (Clayton and Dudley 2005).

Table 1. Material alternatives for transceiver components

Key Component	Material	
Light Source (Laser)	InP/GaAs,	Transmitter
Modulator	InP/GaAs, Si	Transmitter
Laser Driver	Si	Transmitter
Modulator Driver	InP/GaAs, Ge	Transmitter
Mux/Demux, Waveguide	InP, Si	Transmitter / Receiver
PIN/APD detector	InP/GaAs, Ge, SiGe	Receiver
TIA	InP/GaAs, Si	Receiver

Therefore, material platform selection is highly dependent upon a material's optical property, device performance, and potential for component integration. At the extremes, there are two 100GE transceiver designs involving advanced integration:

1) An monolithically integrated photonic circuit in InP
2) A monolithically integrated photonic circuit in silicon containing both active and passive optical components, using processes available in existing CMOS fab.

These advanced designs are under development in university labs and by a small number of start-up firms. Established players in this industry are focusing more on intermediate hybrid approaches for the near term.

2.2.1.1 III-V Group

Material candidates for integration in the III-V group are GaN, GaAs, and InP. (Clayton and Dudley 2005). Several devices on an InP/GaAs platform have already demonstrated success at various degrees of integration. These devices are:

-Tunable lasers and Mach-Zehnder modulators

-DFB laser and Electro-absorption modulators

-Photodetector diodes and TIAs

-Waveguide mux/demux and laser, modulator

-Waveguide mux/demux and detectors

Despite demonstrated feasibility of integration on the InP/GaAs platform, the current high manufacturing cost of integrated InP devices is an obstacle for mass production in the computing, entertainment, and storage markets. Currently, there are few InP fab capable of high volume manufacturing of integrated optoelectronic chips (Clayton and Dudley 2005). Major obstacles include low process yield, small wafer size (50 mm, 2 inch wafer is still common), and liberating engineers from the production line. Nonetheless, if III-V technologies continue to demonstrate superior lasing performance over Si, integration in InP will need to be advanced in order to be combined with the necessary electronic components that can be done in CMOS. There have been some progresses in monolithic integrating III-V/Si. The Compound Semiconductor Materials on Silicon (COSMOS) project at MIT has shown early promises in embedding III-V active element (LED) on a Silicon-on-Lattice-Engineering-Silicon, which is a substrate designed for integrating III-V with Si CMOS (E.A. (Gene) Fitzgerald 2007).

2.2.1.2 Silicon

Established players in the computing industry believe that silicon photonics could provide cost advantages over III-V technologies by leveraging their existing CMOS infrastructure. The Si IC industry is mature and its manufacturing process is the epitome of a convergence between "technological sophistication and economies of scale" (Jalali and Fathpour 2006). The cost of fabricating Si photonics can be reduced by achieving high yield with CMOS compatible, mature manufacturing processes capable of handling large wafer size (200 to 300 mm). Silicon-on-insulator (SOI) wafers could be an ideal platform to fabricate planar waveguide circuit. Key optical elements to be made in Si are modulator, light guiding components, and detectors. Much

research is currently being done on processing Ge on Si platforms for building an integrated receiver, a Si plasma modulator, and even a Ge based laser (Bautista, Morse et al. 2005). Today, monolithic integration of photo-detectors, mux/demux, tapers, and modulators, directly inserted at the CMOS gate level is an area of rigorous research in university labs and private firms. Recent photonic research even demonstrated Si's potentials for optical amplification, lasing, and wavelength conversion (Jalali and Fathpour 2006).

2.2.2 Competing Design

Numerous designs have been proposed for 100GE LAN transceivers. Two key design elements are system architecture and extent of integration. This sub-section presents several design choices for each that are widely discussed in industry.

2.2.2.1 System Architecture

The current industry debates focus on what technology choice would provide the best performance at lowest cost for each transceiver component (Table 2). For each component, there is a variety of available technologies and arrangements to form a 100GE transceiver. For example, there are at least five types of lasers and each can be made to function at 10G, 20G, 25G, and 100G. Finding the appropriate technology and architecture match is a major challenge. Table 3 lists three configurations for 4 x 25G, 5 x 20G, and 10 x 10G architecture proposed by Advanced Photonics Integrated Circuits (APIC) Corporation.

Table 2. Technology-Architecture alternatives

Component	Technology	Architecture
Light Source (Laser)	DFB, DML, EML, VCSEL, Edge Emitter	1. Direct modulation or continuous wave 2. On-chip or off-chip 3. Discrete TOSA or array 4. 10 x 10G, 4 x 25G, 5 x 20G, or 1 x 100G.
Modulator	EAM, MZI, ring	
Mux/Demux, Waveguide, Filters	AWG, Reflective echelle grating, ring resonator filters, thin film filters, splitters, combiners	
Receiver	Surface PD/TWPD	ROSA, flip chip, or monolithic integration

Table 3. Technology-Architecture match proposed by APIC (Khodja 2007)

Architecture / Function	4x25 Gbps	5x20 Gbps	10x10 Gbps
Laser Array	DFB/DBR 1300/1550 nm	DFB/DBR 1300/1550 nm	DFB/DBR 1550 nm
Modulator Array	External Modulation EAM/MZI	External Modulation EAM/MZI	DML/ EAM/MZI
Mux/Demux	Combiner/Interleave	Combiner/Interleave/AWG	AWG
PD Array	Traveling Wave PD	Traveling Wave PD	PD/APD
CWDM/WDM	CWDM	CWDM	WDM, 200GHz

The 4x25G architecture is becoming a popular solution to implement 100G due to its compatibility with the previous 10GBASE-LX4 standard for 10GE. Its main contender, 10x10G, is also attractive, but it may be at a cost disadvantage due to lower yield on fabricating a 10 by 10 monolithically integrated laser array.

2.2.2.2 Integration Scheme

There are a myriad of approaches to integrate key components for a 100GE transceiver. Possible integration scheme includes[6] :

-Laser and modulator

-Modulator, Mux/Demux, waveguide, and detector.

-Mux/Demux, detector, and TIA

- Modulator, Mux/Demux, waveguide, detector, TIA, and driver

Figure 6 displays these schemes in a graphical way. Cells with the same color mean the labeled components below are monolithically integrated.

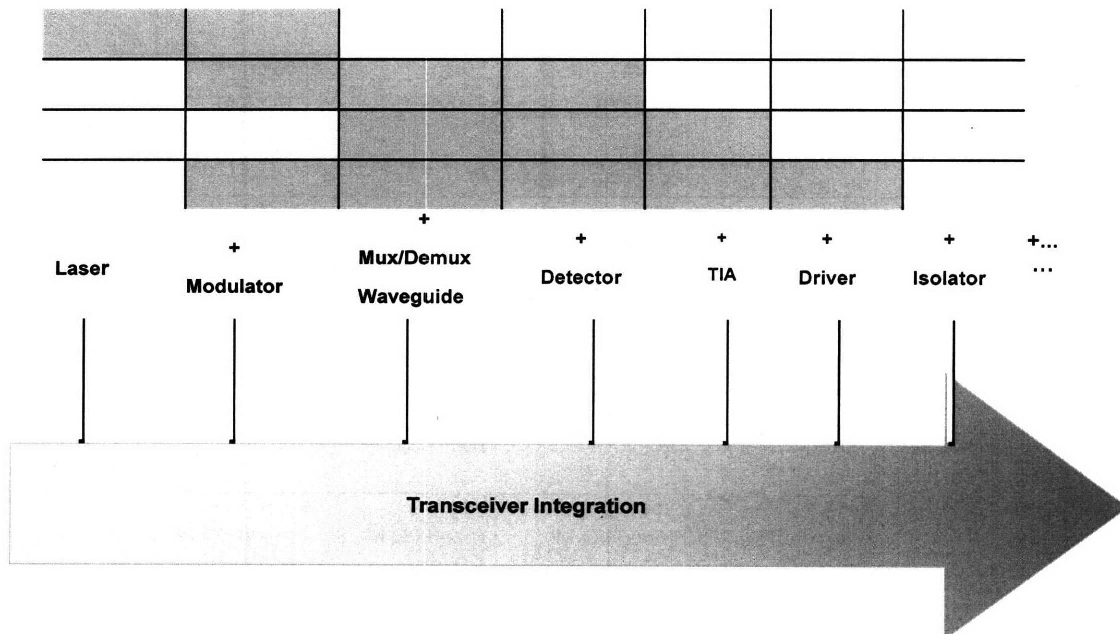


Figure 6. Possible levels of integration for 100GE transceivers

Companies offer and/or have proposed different 100GE solutions ranging from discrete to highly integrated designs in the market. Finisar and CyOptics initially considered discrete designs (Cole 2007). The most discrete solution would be a separately packaged transmitter and receiver. In this case, the TOSA transmitter would be created using multiple TO-CANs. The

[6] Source: Clayton, R. and T. Dudley (2005). Microphotonics: Hardware for the Information Age—Integration in III-V Materials, MIT Microphotonics Center Industry Consortium, Jalali, B. and S. Fathpour (2006). "Silicon photonics." Journal of Lightwave Technology **24**(12): 4600-4615., and O'Brien, D. and M. Schabel (2005). Microphotonics: Hardware for the Information Age: Next Generation Transceivers. Cambridge, The Microphotonics Center at MIT.

ROSA receiver would be made of PIN diodes, combined with an AWG or PLC. (See Table 4 and Table 5)

Table 4. Discrete 100GE optical solution one (Hartman 2007)

Link	*TOSA*	*ROSA*
≤1km	*1.3um 10Ch 10Gb 1dBm*	*PIN w PLC*
5.8dB	*1.3um 5Ch 20Gb 4dBm*	*PIN w AWG or PLC*
10km	*1.3um 10Ch 11Gb 6dBm*	*PIN w PLC*
9.4dB	*1.3um 5Ch 22Gb 8dBm*	*PIN w AWG or PLC*
40km	*1.5um 10 Ch 11Gb 4dBm*	*PIN w PLC*
22dB	*1.3um 5 Ch 22Gb 3dBm*	*PIN SOA AWG PLC*

Table 5. Discrete 100G optical solution two (Cole 2007)

SMF	10km 1310nm	40km 1310nm	10km 1550nm	40km 1550nm
10x10G DML	yes (10λ span needs semi-cooling)	yes (need new DML & RX APD/SOA)	yes (need new DML)	maybe (need new DML)
10x10G EML	yes	yes (need RX APD/ SOA)	yes	yes
5x20G / 4x25G DML	yes (need new DML)	maybe (need new DML & RX SOA)	maybe (need new DML)	no
5x20G / 4x25G EML	yes (need new EML)	yes (need new EML & RX SOA)	yes	yes (need RX DC)
2x50G DQPSK DML	no	no	no	no
2x50G DQPSK EML	yes (need I/Q ML)	yes (need I/Q ML, RX DC & SOA)	yes (need I/Q ML & RX DC)	yes (need I/Q ML & RX DC)

Photonics integrated circuits (PIC) are demonstrated in InP by Infinera (Figure 7) (Jaeger and Perkins 2007). As seen in Figure 7, Infinera offers a separately packaged transmitter and receiver, each containing a highly integrated PIC in InP. A DWDM scheme is achieved by integrating 10 by 10G lasers and modulators with mux in the transmitter, and 10 by 10G detectors with demux in the receiver. On the other hand, Luxtera and Kotura are considering hybrid solutions in Si photonics (Clairardin 2007). An integrated 100GE transceiver using CWDM technology has been proposed in the following scheme (Figure 8). The laser array and detectors are bounded on top of an optical die, which would be fabricated in a CMOS compatible process.

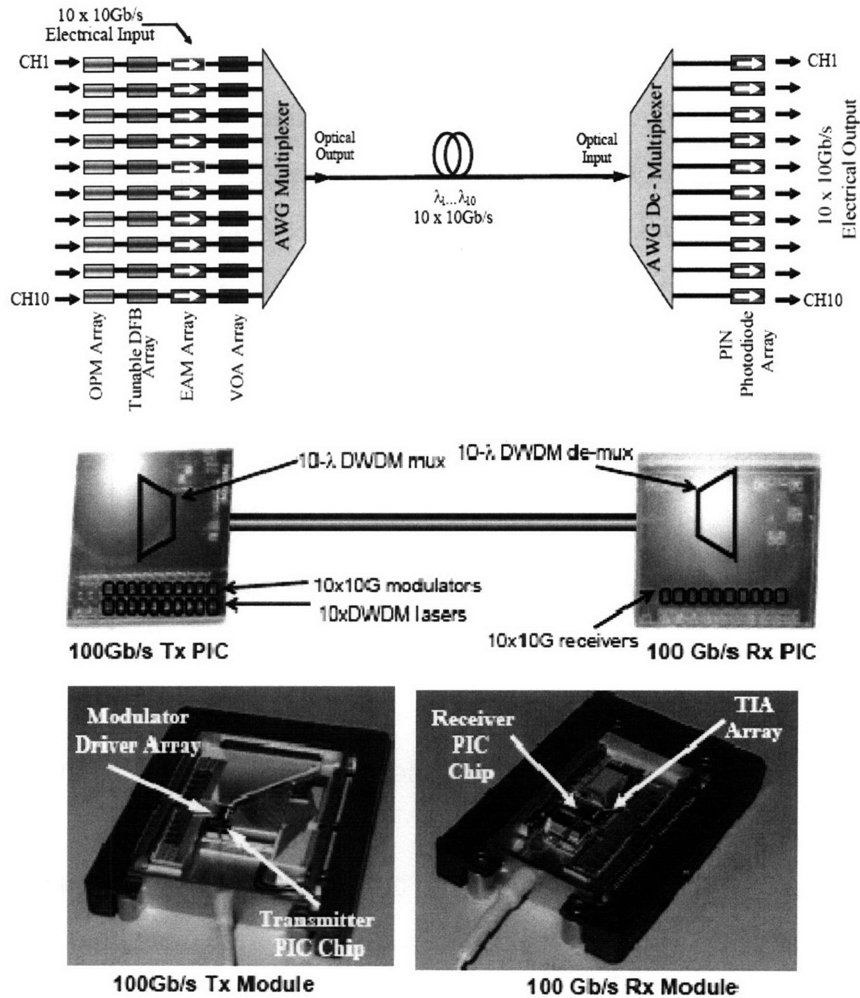


Figure 7. a) A 10x10G DWDM solution, b) Photonic Integrated Circuit, transmit and receive chips, c) Discretely packaged 100GE transmitter and receiver (Jaeger and Perkins 2007)

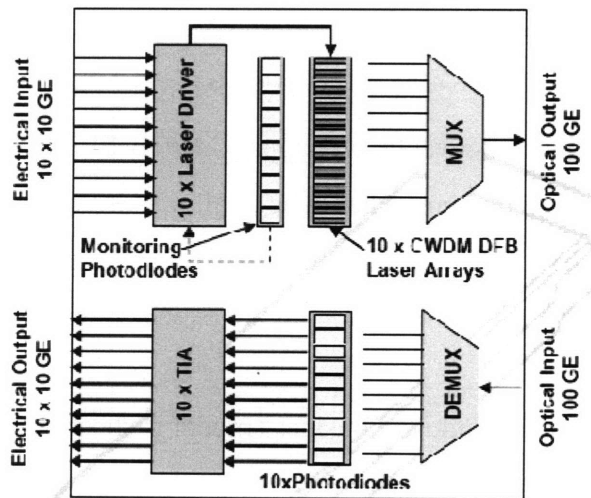


Figure 8. Photonic Integrated Circuit, a CWDM solution—100GE transmitter in a single package (Clairardin 2007)

To summarize, this section shows a variety of designs for a 100GE transceiver that have been proposed and are at early stages of research and development. Selecting the right material platform and design (system architecture and level of integration) among them presents a significant challenge for any firm. A potentially successful design would demonstrate material compatibility at high degree of integration, and appropriate technology-architecture match for a given technical requirement. Furthermore, economic feasibility will be the determining factor to enable an emerging technology to gain commercial adoption. This research employs the process-based cost modeling method to provide insights in this technology selection process in Chapter 3 Method and Chapter 4 Result.

2.3 Current State of the Optoelectronic Industry and the Road Ahead

The optoelectronic (OE) industry has created significant technical innovations that revolutionized the telecommunication market. To penetrate converging markets in the road ahead, established OE components manufacturers and new-entry firms are facing a new set of opportunities and challenges.

The OE industry is recovering from the dotcom bubble of year 2000. Beginning in year 2004, optimism was back in the global marketplace with an increase in venture capital investments. Global laser diode sales enjoyed a 12% sustainable growth excluding the bubble period (Lebby 2006). Projected worldwide OE market demand showed new opportunities of growth in the next decade due to converging applications in communication (telecom, datacom, cable, storage network, FTTP, chip-to-chip), computing (laptop, tablet, desktop, print, virtual, GPS), and consumer electronics (TV, camera, DVD, lighting, biomedical, cell phone). Global optical networking and components revenue was approximately $17 billion in 2005 with strong growth in metro area network equipments sales. The total transceiver market revenue is expected to reach $4 billion by year 2012, in which Ethernet and fiber channel are becoming the dominant sales drivers. In Datacom, 10G transceiver revenues are expected to reach above $2 billion in 2012 (Lebby 2006).

One proposed OE components industry business model is shown in Table 6. In the platform model, a company owns in house R&D, sales, and marketing, but outsources the majority of its manufacturing to third parties. On the other hand, in the vertical model, a company owns the entire product delivery chain, from R &D, to manufacturing, to sales and marketing of the

product (Schmitt 2006). The OE components market could also be divided into high-end and low-end market segments. High-end usually consists of telecom components that have high performance requirements but with low volume (e.g. tunable lasers, 300 pin MSA's). Low end consists of enterprise components that have much higher volume but lower performance variability (e.g. gigabit Ethernet, 10G XFPs).

Table 6. OE components industry business model (Schmitt 2006)

Model	Platform	Vertical
Manufacturing	Outsourced	In House
Example companies	JDSU, Opnext, Avanex, Intel	Bookham, Neophotonics, Avago, Finisar, Luminent

Common beliefs from the semiconductor industry support a vertical model for manufacturing low-end products because vertically integrated companies can take better advantages of economies of scale than companies that outsource their manufacturing. The best model for the OE components industry is less clear. Both of its high-end and low-end market volumes are quite low comparing to the semiconductor industry. Therefore, the platform model may benefit both markets in combining low volume segments to reach economies of scale in production. In both cases, standard, industry coordination, and availabilities of third party foundries are beneficial to both high and low end markets.

Advancements in integration technology have massive implications for the traditional III-V components manufacturers. New Silicon Valley startup firms generated a lot of excitement since their introduction of highly integrated InP transceiver designs in recent years. For example, a new market entry—Infinera built a significant competitive advantage by demonstrating the feasibility of a highly integrated 100G transceiver in III-V materials (Schmitt 2007). Si photonic startups such as Kotura and Luxtera have also demonstrated highly integrated Si photonic based transceivers in 10G and are moving toward 40G to 100G products. In addition, the dominant computing chip maker, Intel, also designs and manufactures optical chips for other companies. Intel started to move into communication chips in 1999 and acquired 36 companies for $11 billion (Kanellos 2002). Intel's optical division may become a full-service outsourcing center for other companies and an intellectual property licenser. The company's advantage is vested in its human capital with extensive knowledge in silicon, and in house manufacturing facilities that can be recycled for photonic production (Kanellos 2002).

In summary, the competitive landscape of the OE components industry is fierce. It is characterized by sophisticated technologies, diverse performance requirements, fast paced product development cycles, and high rates of price erosion (Schabel, Fuchs et al. 2005). The next section of the thesis further emphasizes the importance of standards and coordination for the optoelectronic industry. It discusses roadmapping in integrated photonics, and compares it with success stories from the semiconductor industry.

2.4 The Importance of Standards and Coordination

2.4.1 Introduction to Roadmapping

Roadmapping is a systematic approach to plan for the future. It is a useful tool to encourage technology trajectories and industry coordination. According to a technology roadmapping tutorial written by scholars at MIT, the technology scope of roadmapping can be divided into exploratory mapping and target-drive mapping. The former method is suited for exploration of emerging and potentially disruptive technologies, and the later is more suited for setting specific technical targets for clearly identified technology trajectories. Similarly, the participation scope can be identified as single-organization and multi-organization roadmapping. (Bruce and Fine 2007)

The concept of roadmapping can be combined with the S-curve framework to examine technology evolution. Figure 9 shows the relationship between technology maturity and the types of roadmapping methods in a comparison between integrated circuits and integrated photonics. Integrated photonics is at an early stage of development; thus exploratory roadmapping is used to identify critical technologies and manufacturing platforms.

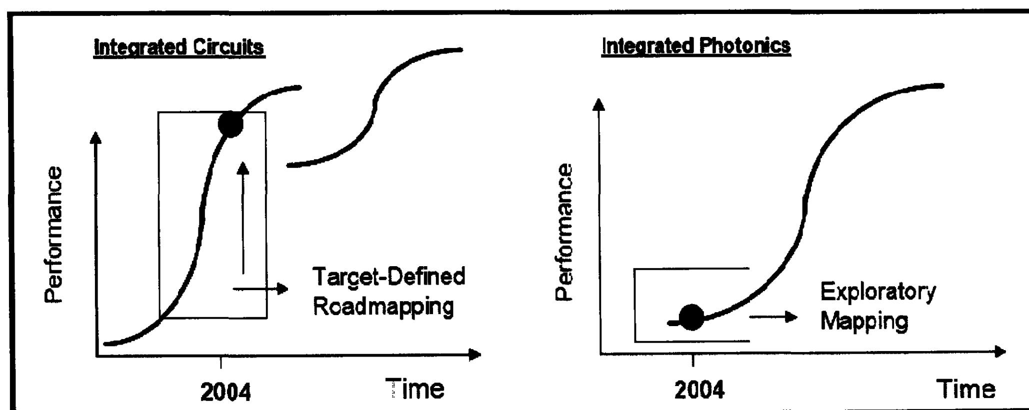


Figure 9. Comparison of electronic and photonic roadmapping on S-curves (Bruce and Fine 2007)

To ensure the health of the optoelectronic industry and draft a cohesive plan for the future, several groups are actively pursuing an industry roadmap. These groups are geographically diverse: in North America, there are the Optoelectronics Industry Development Association (OIDA), National Electronics Manufacturing Initiative (NEMI), groups within the Photonics Manufacturing Association, and the Canadian Photonics Consortium; in Europe, there is the Information Society Technologies (IST) Optimist Program; in Japan, there is the Optical Industry and Technology Development Association (OITDA); and in Singapore, there is the Infocomm Development Authority of Singapore. The International Technology Roadmap for Semiconductors (ITRS), although lacked a photonic roadmap, included an III-V compound semiconductor roadmap in 2003. (Bruce and Fine 2007) In addition, the MIT CTR program was initiated in year 2000 and continuing to provide a fertile ground for academia-industry dialogue of the next generation optical communication systems. The MIT CTR effort can be seen as an exploratory, multi-organizational roadmapping activity.

2.4.2 Lessons from the Semiconductor Industry

Many modern technologies are driven by advancements in microprocessors. As the cost of chip fabrication declines, chips become ubiquitous in many every high-end to mid-end markets. The semiconductor industry fuels the engine of growth in this digital era. It is a major contributor of US economic growth. "The U.S. Federal Reserve Board data shows that while the economy as a whole has grown 30 percent since 1990, the high-tech sector has grown nearly 90 percent" (England and England 1998). In 2006, the semiconductor industry's worldwide revenue stands at $261.4 billion (McCall 2006).

Coordination and standard setting are indispensable in the semiconductor industry. In an industry that lives or dies on the principle of "more for less," demanding ever higher performance and lower cost in a matter of few months, companies across the supply chain quickly learned that cooperation and scale are essential to survival. For example in microprocessor productions, key players are highly coordinated in the realization of each successive technology node generation, guided by Moore's Law[7] as a self-fulfilling prophecy.

---

[7] In 1965, Intel co-founder Gordon Moore predicted that the number of transistors on a chip doubles about every two years.

There are several rationales behind setting industry-wide product and interface standards and coordination on a global scale in the semiconductor industry. Initial investments in R&D and fabrication facilities are large: a state of the art fab has a price tag of $2-3 billion dollars. Standard reduces the risk of investing in the wrong technology. Interoperability and cooperation are essential for high-tech electronic products to function properly. For example, to develop a new microprocessor, engineers need to know the internal workings of PCs, servers, and their operating systems. Scale and efficiency are enhanced through standard, which are necessary in high volume production (x86 microprocessors' global sale is approximately 200 million unit/year). In addition, long term vision and planning allow equipment suppliers to know what technology to expect at what time, and they act years ahead to achieve these goals. Joined technology development enables chip makers to share resources, transfer knowledge, and avoid manufacturing pitfalls. Furthermore, standards enable a platform model— a new industry structure that created the semiconductor foundries and fabless chip design house. This model lowered the barrier of entry and increased innovation and competition.

Semiconductor standards are achieved through highly coordinated international organizations. Two interrelated organizations both create and realize the vision for this industry. The first organization is the International Technology Roadmap for Semiconductors (ITRS)[8], which assesses the industry's technology requirement over the next 15 years to ensure continuous improvement of the integrated circuits and the continuation of Moore's Law. More specifically, the Roadmap identifies a series of technology nodes[9] and their expected arrival dates, usually run on a 2 to 3 years cycle. Currently led by Intel Fellow Paolo Gargini, ITRS is a cooperative effort of global industry manufacturers and suppliers, government organization, consortia, and universities, sponsored by the five leading chip manufacturing regions in the world: Europe, Japan, Korea, Taiwan, and the United States. By region, the US participants are 55% of the total. By affiliation, the chipmakers are 53% of the total participants (Intel 2008).

If ITRS creates long term visions for the industry, the SEmiconductor MAnufacturing TECHnology (SEMATECH) brings these visions to realities. According to Intel, SEMATECH is

---

[8] The sponsoring organizations for ITRS are the European Semiconductor Industry Association (ESIA), the Japan Electronics and Information Technology Industries Association (JEITA), the Korean Semiconductor Industry Association (KSIA), the Taiwan Semiconductor Industry Association (TSIA), and the United States Semiconductor Industry Association (SIA).

[9] Technology nodes refer to the minimum feature size produced on the wafer. Transistor gate length can be a proxy, but it is often smaller than the node length due to techniques used in chip design and layout

the global communication center for members to work together to produce the ITRS. SEMATECH began as a public-private partnership in 1986 with a goal of strengthens the competitiveness of the US semiconductor industry. In 1988, the consortium were consisted of 14 US based manufacturers and the US government. In 1996, the organization shifted focus from the U.S. semiconductor industry to the larger international semiconductor industry by eliminating matching funds from the US government. By 2007, nearly half of the 16 member companies are non-US corporations.[10]

The dominant computing chip maker—Intel's success is partially attributed to the company's heavy involvement in standard initiatives. The company claims that "Intel pursues the latest technological advances by working with more than 100 standards and industry groups worldwide" (Intel 2008). These standards and industry groups span a wide spectrum, including but not limited to computing and consumer electronics platforms, networking and communications, silicon and semiconductors, and software and web. Intel also participates in general standards setting organizations such as the American National Standards Institute (ANSI), European Committee for Standardization (CEN), and European Committee for Electrotechnical Standardization (CENELEC) (Intel 2008).

The optoelectronic industry can learn a great deal from the semiconductor industry's successful standards and coordination framework. Table 7 shows the major differences between these two industries in their technical performance, material usage, processing capability, and existence of a foundry model. In all categories, the optoelectronic industry is lagging behind. If a high degree of photonic-electronic integration requires these two industries to converge, technological, market, and organizational structure and practices are also expected to converge.

[10] SEMATECH's current membership includes AMD, HP, IBM, Infineon, Intel, Micron, National Semiconductor, NEC, NXP (Philips), Panasonic, Qimonda, Renesas, Samsung, Spansion, TSMC, and Texas Instruments. Source: SEMATECH. (2008). "SEMATECH History." Retrieved December, 2007, from http://www.sematech.org/corporate/history.htm.

Table 7. A Comparison between the Optoelectronic and the Semiconductor industry (Bruce and Fine 2007)

	Optoelectronics	**Semiconductor**
Technical Performance	Undefined Mostly discrete	Moore's Law drives device minimum feature size Highly integrated
Material Usage	High diversity	Predominately Si
Processing Capability	Wafer size: 2 to 4 inch Diverse processing equipments Low volume Low yield	Wafer size: 8 to 12 inch Common equipments High volume High yield
Foundry Model	None	Mature

3 Method

This section includes a brief literature search on the existing technical cost modeling methods, a detailed description of the process-based cost model and its application to optoelectronic interconnect production, as well as major data collection and roadmapping activities conducted for this research.

3.1 Existing Cost Modeling Methods

Techno-economic questions constantly arise in a manufacturing process. It is well recognized that any design and process alteration may have significant implication for production cost. Technical cost modeling is a method to provide strategic insights that allows firms to utilize sophisticated engineering theories and mathematical models to project production cost without extensive prototyping and trial and error experimentation (Kirchain and Field 2000). In addition, technical cost modeling is particularly useful in quantifying the risk of innovative designs and processes. It is a tool to compare technology alternatives using a common platform. Previous work in this field have spanned across multiple industries, including electronic packaging (Sikorski, Krueger et al. 1989), printed circuit board (Field and Ng 1989), and material selection in automobile designs (Kirchain and Field 2000).

The optoelectronic industry lacks a standard cost modeling method and sophisticated tools (Ragona 2001). However, in academia, researchers at the MIT CTR have done an extensive study on the cost-feasibility of emerging interconnect technologies and built a sophisticated technical cost model for the III-V material platform. This study focused on the cost advantage of monolithic integration of a 1550nm DFB laser and an electroabsorptive modulator on an InP Platform. (Fuchs, Bruce et al. 2006). A set of related research examined the cost implication of manufacturing offshore; more specifically, production cost of 10G device technologies is compared between U.S. and low-wage Asian locations (Fuchs and Kirchain 2006). Furthermore, cost analysis was conducted to explore the impact of monolithic integration on optical receiver components in realizing 1.55 micron photo-detector on GaAs and Si (Zhang 2004).

In contrast to the optoelectronic industry, the semiconductor and IC industry use well established cost modeling methods and more advanced modeling tools. The industry was able to achieve a highly efficient automated manufacturing process due to its common cost modeling

standard. Industry standards enable effective communication between the equipment users and equipment vendors, and establishes a framework for process development decisions (Ragona 2001).

This standard cost modeling method is called the Cost of Ownership (COO). SEMI E35 defines COO as the "full cost of embedding, operating, and decommissioning, in a factory and laboratory environment, a system needed to accommodate a required volume" (Ragona 2001). Although COO is most commonly used to account for the total cost of acquiring, maintaining, and operating purchased equipment for semiconductor device fabrication, it is also applied to other industries involving heavy machine operations, such as public utilities. The basic concept of conventional COO is shown in the following equations (Nanez and Iturralde 1995):

$$\text{Cost of Ownership} = \frac{\text{Fixed Costs + Reocurring Costs + Yield Costs}}{\text{Tool Life x Throughput x Composite Yield x Utilization}}$$

Which is simplified to:

$$\text{Cost of Ownership} = \frac{\text{Cost To Produce Wafers}}{\text{Number of Wafers Produced}}$$

There are several commercially available cost of ownership software. The TWO COOL(R) model was commercialized through a joint SEMATECH/Wright Williams &Kelly project in 1994. The IC Cost Model was developed by IC Knowledge—a firm consisted of a group of wafer fabrication technologists and management specialists started in year 2000. SEMATECH also developed a Cost Resource Model (CRM). The CRM takes a process flow from SEMATECH workshops, tool parameters, and fab rates as inputs, and outputs equipment requirements, fab costs, and wafer processing costs. The model assumes global factors such as building cost (construction, occupancy, and space), process cost (yield, silicon substrate cost, depreciation rates, and operation hours), and personnel cost (manager, engineers, operators, technicians, and maintenance). Designs can be modeled by different technology nodes and wafer sizes. After the number of wafer starting per month is specified, the model calculates a total wafer cost and a total capital cost by summing capital depreciation, and cost associated with machine specific throughput, personnel per tool, and consumables. (Wright 2001)

Conventional COO is a powerful modeling tool that helps a company comprehends the resource requirement to purchase, operate, maintain, and dispose of an investment, but it still has limitations in predicting production cost for emerging technologies. COO is based on historical data similar to pure accounting methods. In a mature process, equipments and variable costs are

well understood. However, assuming the same type of machines being used for an emerging process and its operating costs being similar to past costs is not a very good assumption when new product architectures, processes, and materials are the very thing under consideration (Liu and Fuchs 2007). Therefore, COO lacks the granularity necessary for projecting cost of emerging product designs and processes. In comparison, the process-based cost modeling method (PBCM) was specifically developed for this purpose. PBCM is introduced in the next section of the thesis.

3.2 Process-Based Cost Modeling

3.2.1 Conceptual Framework

One common misunderstanding in cost modeling practice is the separation of cost analysis from engineering design. Cost is either simplified as a fixed dollar amount; accounting for the price of resources used in production, or is mystified as a magic number that is too "soft" to be accurately quantified using engineering methods. Such a misunderstanding is exaggerated when product design and cost analysis are conducted in separate departments within a firm. The lack of a common language between engineers, accountants, and technology strategists often results in a mismatch between product performance and cost feasibility. Furthermore, the predominant cost modeling methods in industry are retrospective rather than prospective in nature—costs are estimated by accounting for existing resources. When the central question in a cost analysis is the economic viability of an emerging technology, a speculative rather than a normative assessment is desired to consider the effect of engineering design change on product cost. Therefore, cost can be considered as an emergent property that is dependent on product design and manufacturing, as well as other market and organizational factors. (Field, Kirchain et al. 2007)

PBCM is the method chosen for analyzing the cost competitiveness of emerging 100G transceivers in this research because this method factors potential engineering design change into cost analysis. Conceptually, the model can be divided into three functional blocks that incorporate basic engineering principles and up-to-date industry data (Figure 10). The Process Model uses engineering and scientific principles to calculate required processing conditions. For example, in semiconductor processing, deposition rate governed by material properties in a furnace would determine the cycle time of such a process. These results are used to determine the required operating time to meet specified production targets, which translates to the number of required production lines based on the total available operating time in a given year. These results are used to estimate resource requirements—capital, labor, materials, energy, space, etc.—in the Operations Model. Next, the Financial Model maps resource requirements with corresponding operating and investment expenses, and eventually aggregates them into unit cost figures. Ultimately, the model projects the minimum efficient fabrication line that is capable of producing a defined annual volume of good devices and then calculates the cost of installing and operating that line. The scale of the line is determined by the total number of devices (both

acceptable and rejected) that must be processed to achieve the desired annual volume of good units. (Fuchs and Kirchain 2005)

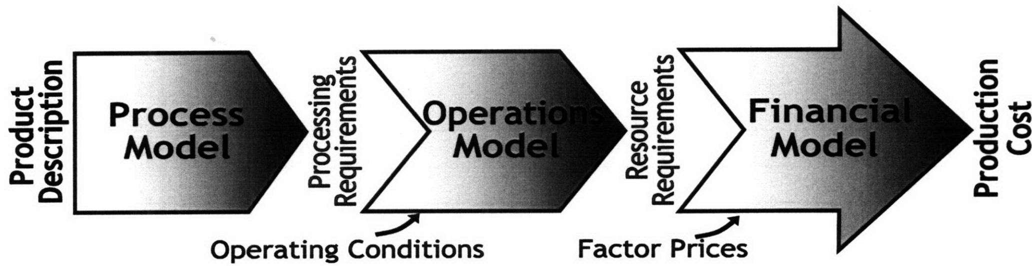


Figure 10. Conceptual framework of PBCM
(MIT Material System Lab)

3.2.2 Model Architecture and Capability

In this study, PBCM is implemented as an Excel based model consisting of multiple worksheets. The top level structure for the Flexible Optoelectronics Production (FOP) PBCM is shown in Figure 11.

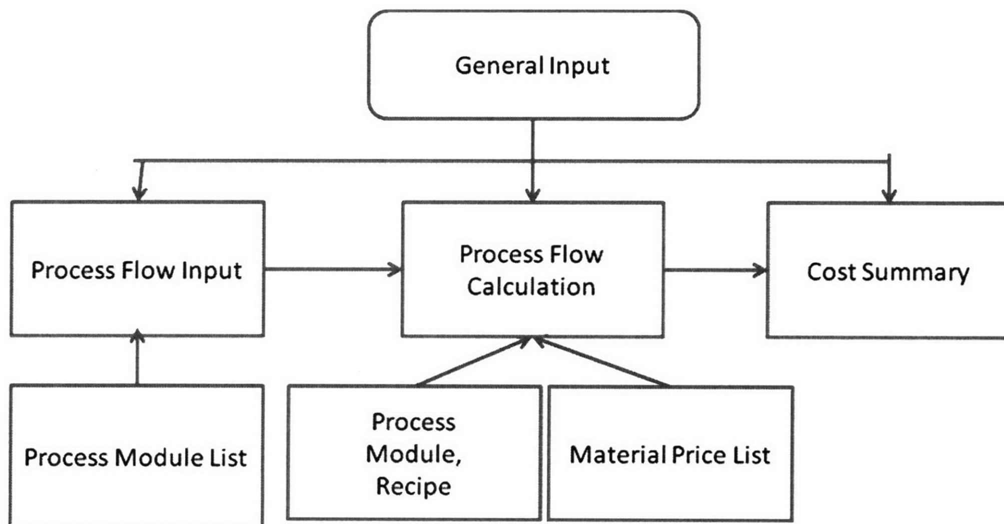


Figure 11. FOP-PBCM Structure

a. General Input

The General Input worksheet contains two top level inputs that drive the model, "Annual Production Volume," which is the number of desired good devices produced per year, and "Product/Process Flow Scenario ID," which indicates the process flow of the product under

analysis. The user can also specify other high level operational parameters for the model shown in Table 8.

Table 8. High level operational parameters on general input sheet

Facility Description	Plant
Working Days/Yr	240
No shifts (hrs/day)	7
Worker unpaid breaks (hrs/day)	1
Worker paid breaks (hrs/day)	1.2
Price Building Space ($ /sq ft) — High Grade Cleanroom	$3,000
Low Grade Cleanroom	$2,000
Non-cleanroom	$1,000
Building Maintenance (% fc)	5%
Building factor	4
Facilities Utilization (Optimized -1, No Addl Capital Expense -2, No Overtime -3, At Cap. -4)	
Capital Dedication (No Override = 0, All Ded = 1, Equipment Sharing = 2, All Non-Ded = 3)	2
MAX_SHARE	90%
Labor	
Ph.D. Wage ($/hr)	$40
Tech Wage ($/hr)	$25
Skilled Wage ($/hr)	$20
Unskilled Wage ($/hr)	$15
Indirect workers/Direct Workers	0.25
Indirect workers/Line	1
General Financial	
Product Life (yrs)	3
Discount Rate	20%
Equipment life (yrs) (default)	5
Installation Cost (%fc) (default)	10%
Maintenance Cost(%fc) (default)	15%
Building Recovery Life (yrs)	25
Price of Electricity ($/kWh)	$0.08
Overhead Burden (% fc)	30%
Design Related	
No. of On-Line Designers	1
Product Designer Salary ($/yr)	$200,000
Pre-Production Product Development Investment	1,500,000

The model can support multiple plant scenarios for both high cost (domestic) and low cost (offshored) manufacturing conditions. These conditions differ in working days per year, number of shifts and breaks, wages, building discount, etc.

b. Process Flow Input

The Process Flow Input worksheet is where the user can manually enter a product's process flow. This worksheet currently contains 39 individual process flows of various designs. For this research, the author created 13 process flows listed below in Table 9:

Table 9. Process Flow Input Table

Name	Material	Wafer size	Wafer Cost	Comp/ Wafer	Comp/ Bar
100G Reflective Echelle Grating-Optical Chip Packaging					
100G Si Two-Chip solution-Optical Chip Packaging					
100G Reflective Echelle Grating (12 inch)	SOI	300	800	245	
100G Si Two-Chip solution Optical Chip (12 inch)	SOI	300	800	210	
Discrete DML	InP	50	250	15000	64
Monolithically Integrated DML Array (10 by 10G)	InP	50	250	7500	32
10G DML TOSA					
100G DML TO-CAN Assembly					
100G Reflective Echelle Grating (6 inch)	SOI	150	150	53	
10G DML TOSA without isolator					
100G Si Two-Chip solution Optical Chip (8 inch)	SOI	200	250	90	
100G Reflective Echelle Grating (8 inch)	SOI	200	250	105	
Monolithically Integrated DML Array (10 by 10G) TOSA					

For each of the products listed under Name, there is a corresponding process flow on this worksheet. The sizes of these flows range from 5 to 119 process steps. In this table, the user also inputs information on material, wafer size, substrate wafer cost, components per wafer, and components per bar, as appropriate, for each product of interest.

b1. Process Module List

When a new process step is entered into a process flow in the Process Flow Input worksheet, first it must be named in the Process Module List worksheet and assigned a category. Available categories include Growth/Deposition, Lithography, Other Front Processes, Assembly Backend, Package Backend, Test, and Optical Sub-Assembly. Currently, there are 68 distinct process steps. These process steps are called Process Modules.

b2. Process Module and Recipe

Process Modules are named in the Process Module List and are created as individual worksheets in the model. Each Process Module can contain multiple sub-modules, called Process Recipes, which are variants of the same process. For example, for the Clean process, recipes include incoming wafer clean, post-lapping clean, etc. Each Process Recipe is described by the

25 input parameters shown in Table 10. These inputs are entered on each Process Module's sheet for each Process Recipe.

Table 10. Process Recipe Input
(Fuchs, Bruce et al. 2006)

Process: (e.g., MOCVD)		
Incidental Yield	Direct Labor: Higher Ed.	Operating Time Per Batch
Embedded Yield	Direct Labor: Technician	Setup Time Per Batch
Machine Cost	Direct Labor: Skilled	Maintenance Freq. (/batch)
Capital Dedication (Y/N)	Direct Labor: Unskilled	Maintenance Time
Capital Usage Life	Installation Cost (%)	Tool/Mask Initial Investment
Max. Batch Size	Maintenance Cost (%)	Tool/Mask Add'l Unit Cost
Average Batch Size	Auxiliary Equipment (%)	High-Grade Cleanroom Space
Unplanned Downtime	Energy Consumption (kWh)	Low-Grade Cleanroom Space
		Non-Cleanroom Space

In summary, the sequence for creating a process flow is 1) create Process Module worksheets containing the appropriate Process Recipes defined with 25 inputs; 2) list the new Process Modules in the Process Module List and assign them categories; 3) input desired Process Recipes in a process flow on the Process Flow Input worksheet.

c. Calculation[11]

The underlying equation for PBCM is: Cost per good device =

$$\frac{\text{Annual Cost of (material + substrate + labor + energy + equipment + tooling + building + maintenance + overhead)}}{\text{Annual good device produced}}$$

Each cost element is aggregated from per-process costs. The Calculation worksheet contains all the per-process cost calculations based on data from the Process Module and Recipe worksheets and Material Price List.

Materials

Materials include both direct materials (those used in building the product) and consumables (chemicals and gases.) For each Process Recipe, the user can either enter a consumption rate for a particular material, or an aggregated rate for general consumables. Materials costs are calculated using the following formula:

$effPV_i = effPV_{i+1}/Y_i \quad i \in [1,\ldots, n-1]$

[11] Equations listed in this section are modified from Fuchs, E. R. H., E. J. Bruce, et al. (2006). "Process-based cost Modeling of photonics manufacture: The cost competitiveness of monolithic integration of a 1550-nm DFB laser and an electroabsorptive modulator on an InP platform." Journal of Lightwave Technology 24(8): 3175-3186.

$effAB_i = effPV_i / Batch_i$

$AC_{Material} = \sum_{i,m} U_i^m * effAB_i * P^m$

effPV: effective annual production volume—gross number of units processed
i: process step number
n: total number of process steps
Y: yield
effAB: effective annual batch
Batch: mean batch size
AC: annual cost
m: material type
U: unit usage of material per batch
P: unit price of material

Substrate

Substrate is the raw wafer base upon which the photonic structure is built. The annual cost of substrate is computed using the following equation:

$AC_{substrate} = Wafer * effPV_n$

Wafer: raw wafer unit cost (dollar per substrate)
$effPV_n$: effective annual wafer volume—gross number of substrates processed to achieve the desired number of good units per year

Labor

Labor is classified as direct PhD, Tech, Skilled, and Unskilled labor. In the PBCM method, labor is not paid based on a fixed annual salary but instead on an hourly wage basis.

The annual cost of labor is calculated as follows:

$AC_{labor} = \sum_{i,l} APT_i^l * P^l$

$APT_i^l = DPY * (24 - NS - UB) * WPL_i^l * LR_i$

$LR_i = reqLT_i / availLT$

$reqLT_i = effAB_i * cycleT$

$availLT = DPY * (24 - NS - UB - PB - UD)$

APT: annual paid labor time
l: labor type (direct PhD, Tech, Skilled, and Unskilled)
DPY: operating days per year
NS: no operation—closed hours per day
UB: unpaid breaks (hours per day)
WPL_i^l: fractional labor type l assigned to step i
LR: lines required

reqLT: required operating time
availLT: available operating time
cycleT: cycle time
PB: paid breaks (hours per day)
UD: unplanned downtime (hours per day)

Figure 12 displays the manner in which the uptime and downtime of a manufacturing line is conceptualized in the model.

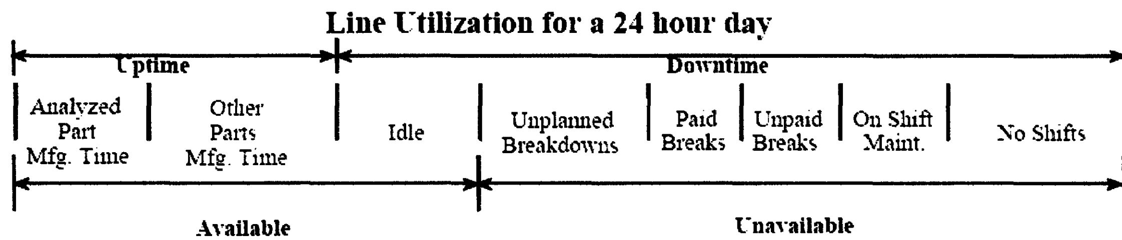


Figure 12. Category of time breakdown based on line utilization for a 24 hour day (Fuchs, Bruce et al. 2006)

Energy

Energy costs are calculated based on user's input of power consumptions per process step. The formula is:

$AC_{Energy} = \sum_i reqLT_i .* Po_i$

Po: power usage

Capital Cost (Equipment, Tools, Space)

In the FOP-PBCM, the cost of capital resources is calculated using the capital recovery factor, which distributes initial investment of capital uniformly over its usage life. The discount rate represents the opportunity cost associated with laying down costly investment upfront. The cost of equipment, tools, and space is computed using the following formula:

$CRF_x = \frac{r(1+r)^{t_x}}{(1+r)^{t_x}-1}$

$AC_x = I_x * CRF_x$

CRF: capital recover factor
r: discount rate
x: equipment, tools, space
t: time period over which investment is distributed. Equipment, tool, and building can have different time periods.
I: the up-front investment to be allocated

Equipment cost is accounted for with special care in the model. A user can specify whether a machine used in a Process Recipe is non-dedicated (shared with other products,) shared (between the same Process Recipes,) or dedicated. These specifications are entered either as global default values on the General Input page, or entered individually for each process. In the non-dedicated scenario, equipment investment is allocated according to the fraction of line time spent in manufacturing the product of interest. In the shared scenario, processes using the same equipment are modeled as if executed on the same machine, subject to a maximum share factor. For example, max_share = 0.95, means if a machine is 95% utilized at step i, the remaining 5% machine time cannot be shared. In the dedicated scenario, the machine cost is allocated to the integer (rounded up) line time at step i. In mathematical expressions:

$R = I * CRF$

$AC_{equipment} = AC_{non\text{-}dedicated} + AC_{shared} + AC_{dedicated}$

$AC_{non\text{-}dedicated} = \sum_i R_i * LR_i \quad i \in \{non\text{-}dedicated\}$

$AC_{shared} = \sum_j R_j(\lceil \sum_i (LR_{ij} - \lfloor LR_{ij} \rfloor)_{max_share} \rceil + \sum_i \lfloor LR_{ij} \rfloor)$

$i \in \{shared\}, j \in \{1, \ldots J\}$

$AC_{dedicated} = \sum_i R_i .* \lceil LR_i \rceil \quad i \in \{dedicated\}$

R: allocated cost for the period t
max_share: a factor denoting the maximum percentage of equipment time to be shared
i: process step
j: Process Recipe

Auxiliary equipment cost is calculated as a percentage of the equipment cost. This percentage factor can be a global input as well as varying across equipment.

$AC_{auxiliary\ equipment} = AC_{equipment} * AE$

AE: auxiliary equipment factor as a percentage of the equipment cost

Maintenance and Overhead

Maintenance and overhead cost are calculated using the formula below. Similarly, MF, BMF, and OBF can be global as well as per-process inputs.

$AC_{maintenance} = (AC_{equipment} + AC_{auxiliary\ equipment}) . MF + AC_{building} * BMF$

$AC_{overhead} = (AC_{equipment} + AC_{auxiliary\ equipment} + AC_{tool} + AC_{building}) * OBF$

MF: maintenance factor in percentage

BMF: building maintenance factor in percentage
OBF: overhead burden factor in percentage

Yield

Common modeling methods in industry often assume a global yield number for the entire process flow. The FOP-PBCM has the capability of assigning two types of yield per process: an incidental yield and an embedded yield. Incidental yield takes account of obvious yield loss at a given step that can be seen without testing, such as physical cracks and scratches. Embedded yield represents non-obvious defects in a step that cannot be discovered until reaching a later test step. Test steps can only have incidental yield. The yield at a test step is calculated by the product of incidental yield at the test step and previous embedded yields right after the most recent prior test step.

$$Y_{i=k} = \begin{cases} incY_k * \prod_{x=(t^*+1)}^{k} embY_x, & k = test \\ incY_k, & k \neq test \end{cases}$$

Y: yield
incY: incidental yield
embY: embedded yield
k: indicator for a given process step
t^* : most recent prior test step to k. If there is no prior test step, start with the first step.

c1. Material Price List

The Material Price List worksheet currently contains prices for 49 types of materials in categories such as General Cleaning, Lithography, Deposition and Epitaxy, and Etching, and 58 types of backend packaging and assembly components. These prices are used by the Calculation worksheet.

d. Cost Summary

The Cost Summary sheet contains the outputs of the FOP-PBCM (Table 11)

Table 11. Final result table of FOP-PBCM in Cost Summary sheet

AGGREGATE PROCESS COST SUMMARY				
VARIABLE COSTS	**$/product**	**$/year**	**percent**	
Material Cost	--	--	--	
Labor Cost	--	--	--	
Energy Cost	--	--	--	
Substrate Wafer Cost	--	--	--	
Total Variable Cost	--	--	--	
FIXED COSTS	**$/product**	**$/year**	**percent**	**investment**
Main Machine Cost	--	--	--	--
Auxiliary Equipment Cost	--	--	--	--
Tooling Cost	--	--	--	--
Building Cost	--	--	--	
Maintenance Cost	--	--	--	
Fixed Overhead Cost	--	--	--	
Total Fixed Cost	--	--	--	
TOTAL FABRICATION COST	--	--	--	
Current Scenario:	--			
Total Investment	--			
Investment. Weighted Utilization	--			

In this table, total fabrication cost for a product of interest is broken down into variable cost (material, labor, energy, and substrate wafer cost), and fixed cost (main machine, auxiliary equipment, tooling, space, maintenance, and fixed overhead) in units of dollars per product and dollars per year. The total investment is the upfront cost of purchasing the main machine, auxiliary equipment, and tools. Investment weighted utilization (IWU) is a ratio of the sum of the percentage of line required over the sum of number of lines allocated (varies by dedicated, non-dedicated, and shared equipment types)—a number to indicate factory utilization rate:

$$IWU = \frac{\sum_i LR_i}{\sum_i \text{Line Allocated}}$$

In addition to representing high level outputs of the model, the Cost Summary worksheet also presents cost breakdown by process category, specified in the Process Module List worksheet (Lithography, Growth/Deposition, Other Front, Test, Package Backend, Assembly Backend, and Optical Sub-Assembly), as well as cost breakdown by Process Module. The numbers of equipment allocated for each Process Module are also shown on this worksheet.

3.3 Building Virtual Manufacturing Facilities

The PBCM method establishes a framework for projecting the resource requirements of a virtual manufacturing facility. PBCM is applied to this research in modeling a real world III-V material based fabrication facility (a/k/a fab) commonly seen in the photonic industry, and a Si material based fabrication facility commonly seen in the semiconductor industry.

3.3.1 III-V Fab

Previous CTR modeling efforts in the III-V space focused on a transmitter—the integration of an InP based laser and modulator. Building on pre-existing Process Modules in the FOP-PBCM, this research created additional modules and recipes listed in Table 12. These processes are used for the TO-Can and DML Array transceiver designs.

Table 12. Process Module and representative Process Recipe for III-V Fab

Surface Treatment	Growth/ Deposition	Etch	Lithography	Thermal	Test
Clean	**MOCVD**	**Plasma Etch**	**HMDS Prime**	**Anneal**	**Post Plant Transfer**
a. Incoming Wafer Clean	a. InGaAsP on InP (varying thickness)	a. Etch SiN (varying thickness)	**Spin-On Resist**		**Inspection**
b. Post-lapping Clean	b. Undoped InP Overgrowth	b. Etch SiO2 (varying thickness)	**Prebake**		**Wafer Incoming Inspection**
c. Die Plasma Clean	c. P-type InP	**Wet Etch**	**Litho**		**Bar**
Spin dry	**Metal Liftoff**	a. Etch SiN (varying thickness)	**Develop**		
Lapping	**PECVD**	b. Etch SiO2 (varying thickness)	**Postbake**		
Wafer Cleave	a. SiN (varying thickness)	c. Grating Etch	**Ebeam**		
Bar Cleave	b. SiO2		**Holography**		
HR Coating					
AR Coating					

3.3.2 Si Fab

One original goal of this research is to create new Si process modules in FOP-PBCM. These modules are used to model the Hybrid and Si Two Chip transceiver designs that will be described in Chapter 4. Table 13 lists these new Si modules, which include both standard semiconductor processes and special photonic processes.

Table 13. Process Module and Process Recipe for Si Fab

Surface Treatment	Growth/ Deposition	Etch	Lithography	Thermal	Test
Clean	**Epi Growth**	**Resist Strip**	**0.25 um**	**Anneal**	**Measure**
a.H2SO4:H2S2O8/ HF/ SC1/SC2	a. Si	a.10:1(H2S04: H2S2O8)	**0.18 um**	**Densification**	a. Film Thickness
b. O3/Anhydrous HF	b. SiGe	b. Active Area Oxy Strip, 15:1 HF	**0.25 waveguide**	**Implant Oxide**	b. CD
CMP	**Ion Implantation**	c. Plasma O2 Ash			c. Overlay
Wafer Scribe Cleave	a. High Energy	**Plasma Etch**			**Inspection**
a. 6 inch	b. Low Energy	a. Etch SiO2			a. Defect
b. 8 inch	**Sputter (Al-Cu)**	b. Si Dry Etch			b. Optical
c. 12 inch	**PECVD**	c. Etch Taper			**Incoming Inspection**
	a. Deposit Undoped HDP Oxide	d. Etch Gate			
	b. Deposit poly-si	e. Etch Metal			

3.3.3 Backend

Backend packaging and assembly process modules are listed in Table 14. These modules are material platform independent. Due to the large number of process recipes associated with some of the modules, individual recipes are not listed in Table 14.

Table 14. Backend Process Module

Assembly	Optical Sub-assembly	Package	Test
Burn In	**Alignment**	**Bake**	**Visual**
Chip Bond	**Adhesive Dispense**	**Lidding**	**Assembly**
Wirebond	**Mount**	**Package Clean**	**Isolated Die**
Filter Assembly	**Weld**	**Fiber Attach**	**Leak**
	Die Polish	**Code and Label**	

3.4 Data Collection

Investigating the economic viability and operational hurdles of emerging transceiver designs requires accurate, up-to-date design and manufacturing details from the optoelectronic industry. Data collection includes initial transceiver designs, process flows for these designs, and operational data—25 input parameters for each Process Recipe (machine cost, cycle time, labor, footprint, yield, etc.) from the component manufacturers.

Previous CTR data collection effort in III-V component modeling involved "accurate-to-industry" details of 20 firms across the optoelectronics supply chain (Fuchs, Bruce et al. 2006). For this research, transceiver designs were obtained from design engineers in industry as well as mined from the literature to identify hotly debated alternatives in the IEEE HSSG interim meetings' publications over a period of one and a half years. The authors collected process flows and operational data through field work, travelling to half a dozen of the world's largest optics and electronic companies across the value chain to conduct on-site interviews and fab tours. For III-V production, operational data were aggregated with data from the previous work mentioned above to achieve an industry average. For Si production, operational data were obtained from publications and from interviews at SEMATECH. Data from SEMATECH are representatives of most industry conditions. After establishing initial contact and on-site visitations, the bulk of the data-gathering effort was conducted through follow-up phone interviews with top-level managers, design engineers, factory line managers, equipment suppliers, sales associates, and cost model builders.

3.5 Roadmapping Method: Cross Market Interviews

This study addresses the embedded policy question on the role of standards and coordination (roadmapping) as enablers of emerging technology in the optoelectronic industry. The goal is to gather market size projection data to provide context to the cost results. A case study of basic concepts of roadmapping and photonic industry roadmapping is introduced in the background Section 2.4. This section focuses on introducing the roadmapping approaches deployed at the MIT Microphotonics Consortium. In particular, a cross market questionnaire is jointly developed and conducted by Jonathan Lindsey, Shan Liu, and Yaoqi Li. Evolution of this questionnaire is explained in this section.

3.5.1 General Roadmapping Methodologies

Multiple approaches to roadmapping activities exist, and often are tailored to the particular industry under consideration. One basic methodology is proposed in Figure 13.

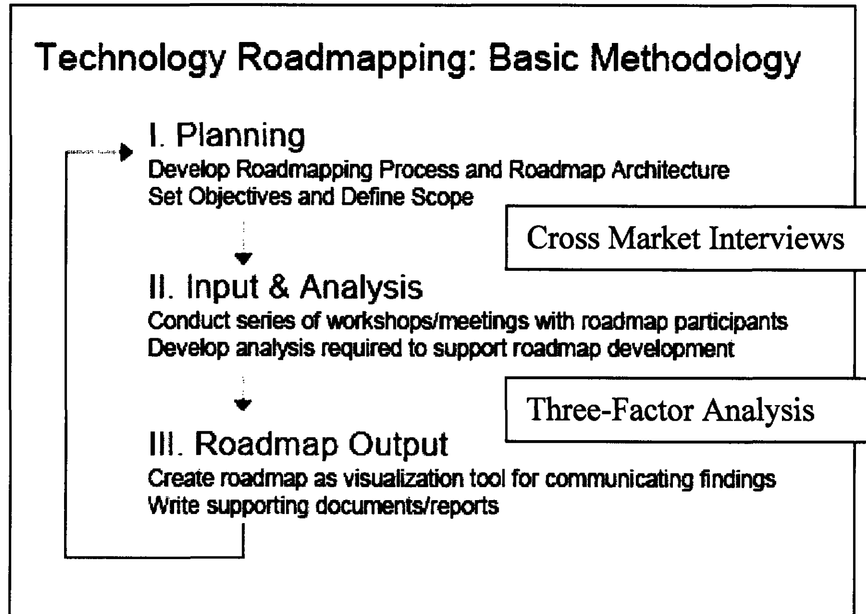


Figure 13. Stages of technology roadmapping
(Bruce and Fine 2007)

The purpose of participating in MIT MPC's roadmapping effort is to seek an understanding of the technical, market, and organizational challenges posed in gaining commercial acceptance of Si photonic technology, and the possible spillover effects of this new understanding from one related industry to another. To assist the consortium's multi-organizational exploratory roadmapping effort, a cross market questionnaire was developed at MSL with the aim of interviewing high-level managers and engineers at various companies across markets. Four major market segments were targeted: computing (chip-to-chip, chip-to-board, board-to-backplane) datacom (high performance computing) video, automotive (media, sensor) and wireless handheld devices. For this research on optical transceiver designs, the relevant markets are computing and datacom. The questionnaire fits nicely into the roadmapping Planning and Input stages as indicated in Figure 13. Most importantly, output of these interviews is visualized as a three-factor analysis, mapping out production cost at each transceiver technology and volume combination, and then comparing this cost with price expectation to determine the

viability of such transceiver markets in the datacom and computing industries. The result of this analysis will be presented in Section 4.4.

3.5.2 Cross Market Questionnaire

Five iterations of the questionnaire were developed over four months. One example of the final version for the computing market is included in Appendix B. One major challenge is the lack of a common language across industries. For example, the system, component, and semiconductor engineers don't use the same vocabulary for the same concepts. Furthermore, even the "same" words often have different meanings. These challenges will be discussed more in detailed at the concluding chapter of this thesis. The interviewer must interpret the interviewees' particular meaning accurately, and remain consistent across interviews. Otherwise, outputs of the questionnaire may be comparing apples with oranges. The final version of the questionnaire explicitly includes all the underlying assumptions for an interviewee's technology/cost projection. For example, asking three questions on data rate per link, number of links per system, and number of links per transceiver allows the interviewer to calculate the annual unit volume of bandwidth equivalent100G transceivers consistently across interviews in spite of different definitions of transceivers. Initial results of the first round of interviews will be presented in Section 4.4.

4 Results

This research investigates the economic feasibility of producing a 1310nm, single mode, CWDM, 100G Ethernet LAN transceiver. As discussed in Section 2.2, firms face strategic decisions on integration and material selection among a variety of 100GE designs. To better understand the impact of integration and material platform on production cost, four designs for a functionally equivalent 100GE LAN transceiver are investigated: (1) TO-CAN: a discretely packaged InP transmitter consisting of 10 TO-CANs and a discretely packaged receiver; (2) DML Array: a discretely packaged InP transmitter consisting of a 10 by 10G directly modulated laser (DML) array and a discretely packaged receiver; (3) Hybrid: a hybrid transceiver consisting of an InP DML array, a III-V detector array, and an integrated Si photonic chip in a single package; (4) Si Two Chip: a hybrid transceiver consisting of an InP DFB laser array and a monolithically integrated Si photonic chip in a single package. The details of these designs are explained in Section 4.1.

The four designs are modeled in the FOP-PBCM using data on existing processes collected from industry. Operational data for emerging processes are estimated from existing processes by experienced engineers. All four designs are modeled using the process modules listed in Table 12, Table 13, and Table 14 in Section 3.2. Cost analyses of the InP and Si designs are presented in Section 4.2 and 4.3. The scope of these analyses is limited to front-end production and back-end assembly of functionally equivalent 100GE optical devices.

4.1 Layout of Four Designs

4.1.1 Design 1. TO-CAN

The transmitter is made of ten discrete10G TO-CANs, individually packaged and then aligned with thin film filters. The hybrid receiver consists of an AWG, 10 photo-detectors, and a TIA. The transmitter and receiver are discretely packaged first, and then combined into an outer package (Figure 14).

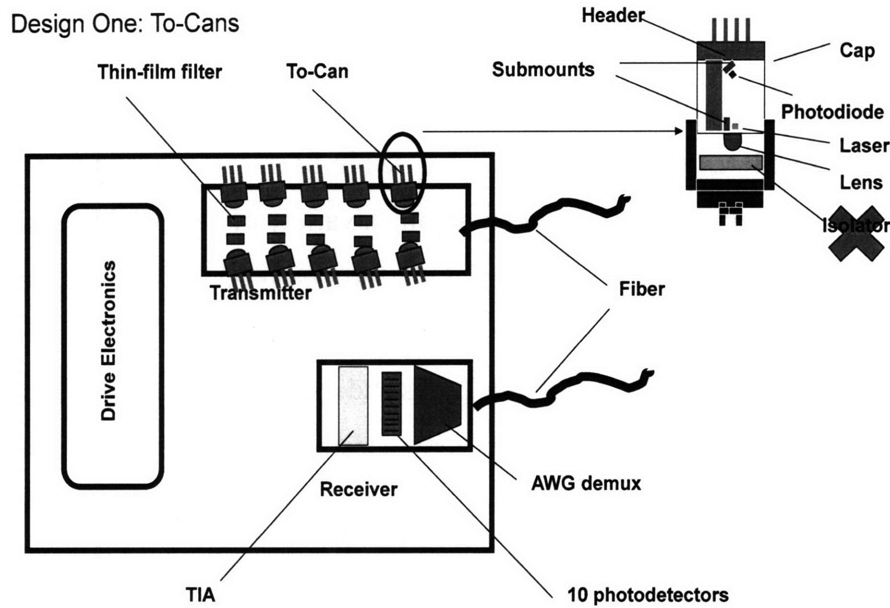


Figure 14. TO-CAN

4.1.2 Design 2. DML Array

The transmitter is made of a DML array (monolithically integrated 10 by 10G lasers in InP), and aligned with thin film filters deposited on triangular prisms. The hybrid receiver consists of an AWG, 10 photo-detectors, and a TIA (same receiver as in Design 1). The transmitter and receiver are discretely packaged first, and then combined into an outer package (Figure 15).

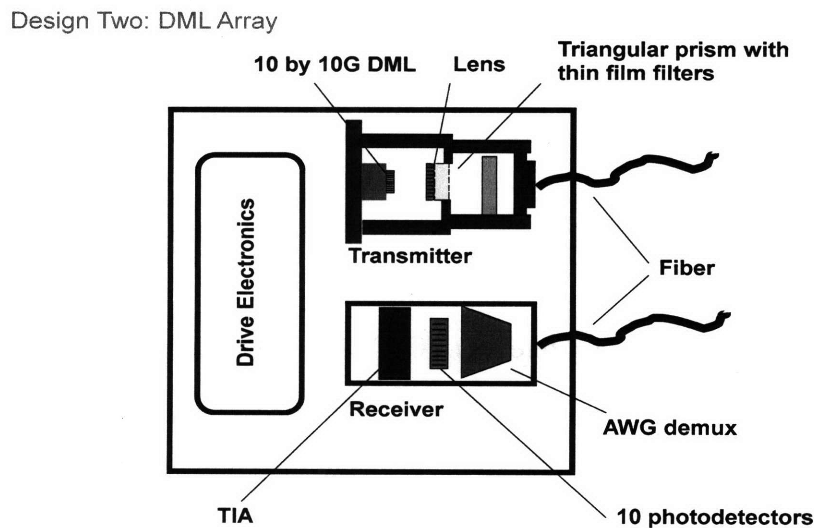


Figure 15. DML Array

4.1.3 Design 3. Hybrid

A hybrid transceiver with a DML array (10 by 10G DML, same as in Design 2), 10 monitoring photo-diodes (MPDs) and 10 photo-detectors made in a III-V material, are all mounted on top of an integrated Si die with waveguide (including mux/demux) (Figure 16).

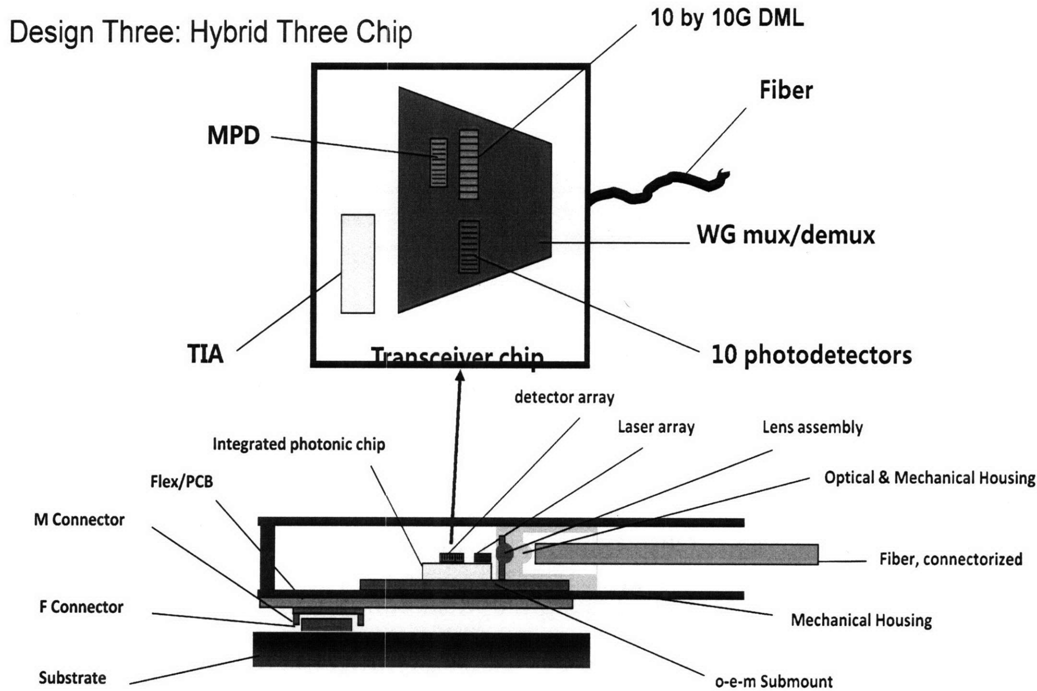


Figure 16. Hybrid

4.1.4 Design 4. Si Two Chip

Design 4 involves advanced integration of photonic functions. The optical chip is a monolithically integrated photonic circuit comprised of waveguide, taper, Si modulator, and Ge photo-detector. The light source is still an InP DFB laser array, flip chipped on top of the Si die. The process flow for this device is currently assumed at the 0.18 micron technology node (Figure 17).

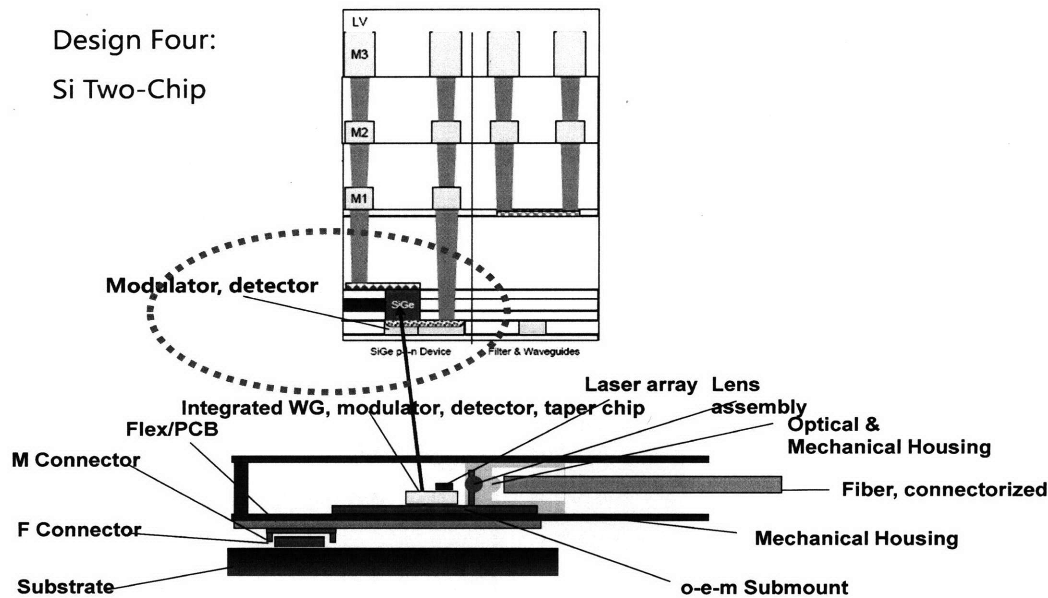


Figure 17. Si Two Chip

4.1.4.1 One Chip Solution

This research also studied a One Chip solution involving advanced integration techniques most suited for the computing space. In this design, optical components are monolithic integrated with CMOS logic. The photonic portion of the device is only microns in size, encompassing all of the required elements except light generation, including waveguide routing, filters, modulators, and detectors. Some researchers believe that this is the only viable approach for serious consideration of photonic-electronic integration (Beals 2007). Microprocessors are on the order of 10 mm by 15mm on a 150 nm (single core) node and shrinking rapidly with each successive node generation. The One Chip solution is most likely to be sized in the same range. Research conducted at MIT started with a design based on a 180 nm node process flow using Al/Cu interconnects technology. Migration of this flow will be to a 150nm node and lower. Practical insertion of photonics is most likely at the 90 nm CMOS node depending on application performance. If 90 nm becomes the insertion point, an interesting application would be supporting communication at even higher levels of electronic integration such as inter-core communications for multi-core processors (Beals 2007). The One Chip solution is outside of the scope for this study, but we propose a lower bond cost estimate of producing such design in Section 4.1.4.1. It is a fascinating area deserving in-depth future research.

4.2 Cost Analysis on Integration

This section provides insights on the cost competitiveness of integrating multiple components on a single device during frontend fabrication, and its implications for backend packaging and assembly. In order to isolate the effect of integration on cost feasibility, analyses are conducted for the InP and Si designs independently. For each set of designs, three analyses are the focus: 1) quantifying the impact of economies of scale, 2) identifying the top cost drivers, and 3) exploring the sensitivity of cost to production yield and which yield improvements lead to significant cost savings.

4.2.1 InP Designs

The InP designs includes Design 1 TO-CAN and Design 2 DML Array. This section presents a cost comparison of these two designs.

4.2.1.1 Frontend (Lasers) Cost Comparison

Frontend production is defined as all processing steps prior to packaging and assembly. It usually starts with the first incoming wafer clean step and continues all the way to the final die inspection step. For the InP designs in this study, frontend is the production of lasers. Process flow for Design 1 has 59 steps and Design 2 has 67 steps. The main difference between the two frontend fabrication processes is that in Design 1, lasers are diced to the 10G functional level; and in Design 2, lasers are diced to the 100G functional level to form the 10 by 10G laser array. Table 15 lists the underlying assumptions in the model:

Table 15. Assumptions in the frontend fabrications of Design 1 and 2

	Design 1. Discrete 10G laser	Design 2. Integrated 100G laser array
Wafer size	2 inch	2 inch
Material	InP	InP
Die/Wafer	15,000 (10G)	7,500 (100G)[12]
Substrate Cost	\$250	\$250
Frontend Yield	3%	0.3%

[12] Each die consists of a 10 by 10G laser array

Notably, the integrated design places five times as many 10G lasers on a wafer compared to the discrete design. Lasers are more densely packed because 1) the technical requirement of array lasers specifies a higher laser density on a single die, and 2) less dicing space is needed to separate die with larger size.

4.2.1.1.1 Economy of Scale

Figure 18 displays the modeled costs for the discrete laser and the DML array. Production of both designs show strong economies of scale up to annual production volumes of approximately 1 million equivalent 100G units. At annual volumes above 1 M units, the production costs of the two devices level out: the ten discrete 10G lasers at just above $24 per unit, and the 100G DML array at just below $43 per unit. This result indicates the integrated device is always more costly than the aggregate of discrete devices on the frontend.

Final product yields of 3% for the discrete laser in Design 1 and 0.26% for the DML array in Design 2 are the worst-case scenarios. Since the laser yields are in the single digits and lower, slight improvements within individual process steps would be expected to lead to significant savings on production cost. A detailed analysis on yield sensitivity is presented in Section 4.2.1.1.3.

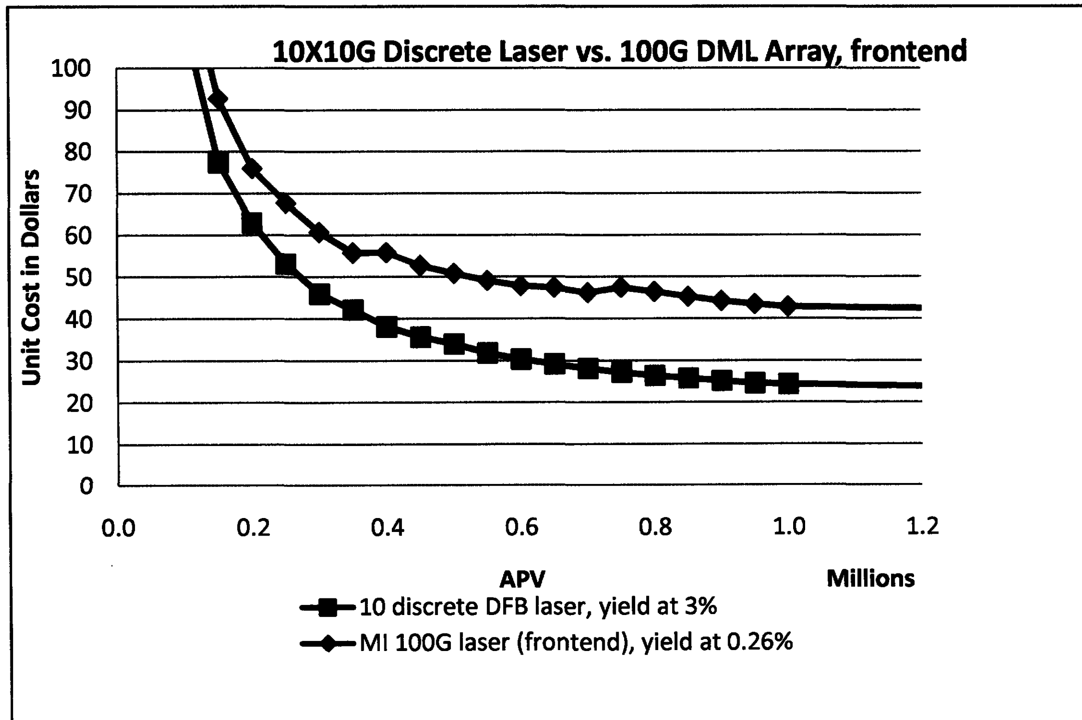


Figure 18. 10 by 10G discrete laser vs. 100G DML array cost comparison (frontend)

4.2.1.1.2 Top Cost Drivers by Process Types and Cost Element

Figure 19 and Figure 20 present the seven processes that represent the largest fraction of the total modeled cost of the discrete laser and DML array production at 1M APV. These figures also show a breakdown of cost elements (materials, other variable costs[13], equipment, and other fixed costs) within each process. Bar cleave, MOCVD, e-beam, and lithography are dominated by equipment and other fixed costs, while HMDS and wet etch are dominated by material costs. Interestingly, at relatively high volumes (>1M APV), labor intensive processes become the dominant cost drivers (i.e., bar test), since fixed costs become less significant as they are amortized across more units. Comparing the two designs, the DML array requires higher expenditure on lithography but lower expenditure on cleaving (breaking wafers to individual dies) compared to discrete lasers. At 1M APV, total allocated investment (equipment and tools) for Design 1 is $31.6 M, and for Design 2 is $48.8 M. The integrated design requires a much higher capital investment due to its low production yield.

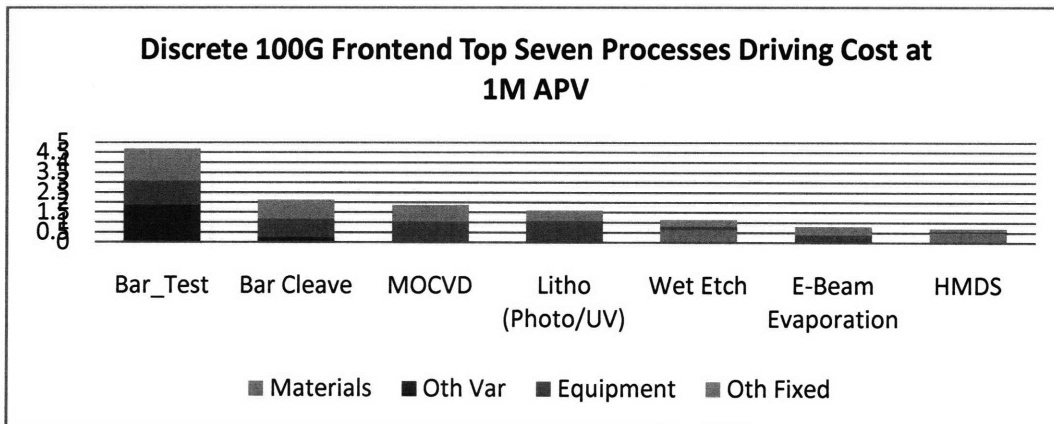


Figure 19. Discrete lasers frontend top seven processes driving cost at 1 million APV
Total Unit Cost $18.8 (excluding substrate cost)

[13] Other Variable Cost includes labor and energy costs. Other Fixed Cost includes tool, building, maintenance, and overhead costs.

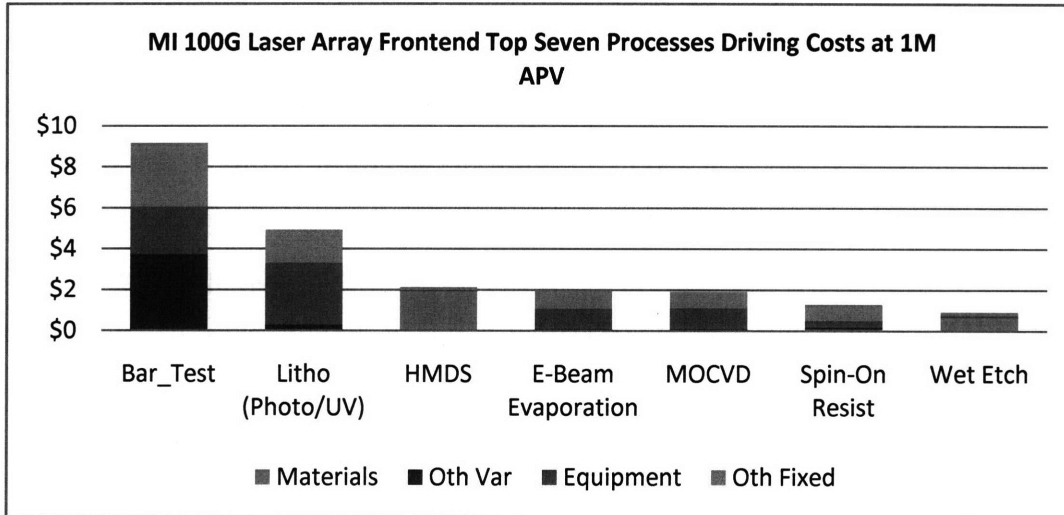


Figure 20. DML array frontend top seven processes driving costs at 1 million APV
Total Unit Cost $30 (excluding substrate cost)

4.2.1.1.3 Sensitivity of Cost Estimate to Production and Process Yield

The impact of yield improvements on the frontend production cost of lasers is quantified in Figure 21 and Figure 22. Figure 21 indicates that due to the inherent cost savings opportunities associated with the integrated array, it is not necessary to reach yield parity to match and drop below the cost of the discrete lasers. Specifically, the figure shows that an increase to 2.61% (admittedly a ten-fold increase from currently estimated conditions) leads to a 50% cost advantage at 1M APV. Figure 22 plots the discrete laser cost against varying frontend yields. At a 10% frontend yield, the cost is cut in half from the cost at a 3% yield. The yields shown in the figures are the aggregated frontend yields, obtained by varying per-process yields at the bar cleave and bar test steps. This type of analysis allows firms to make strategic decisions on the amount of effort to dedicate to overall yield improvement in reaching a cost target.

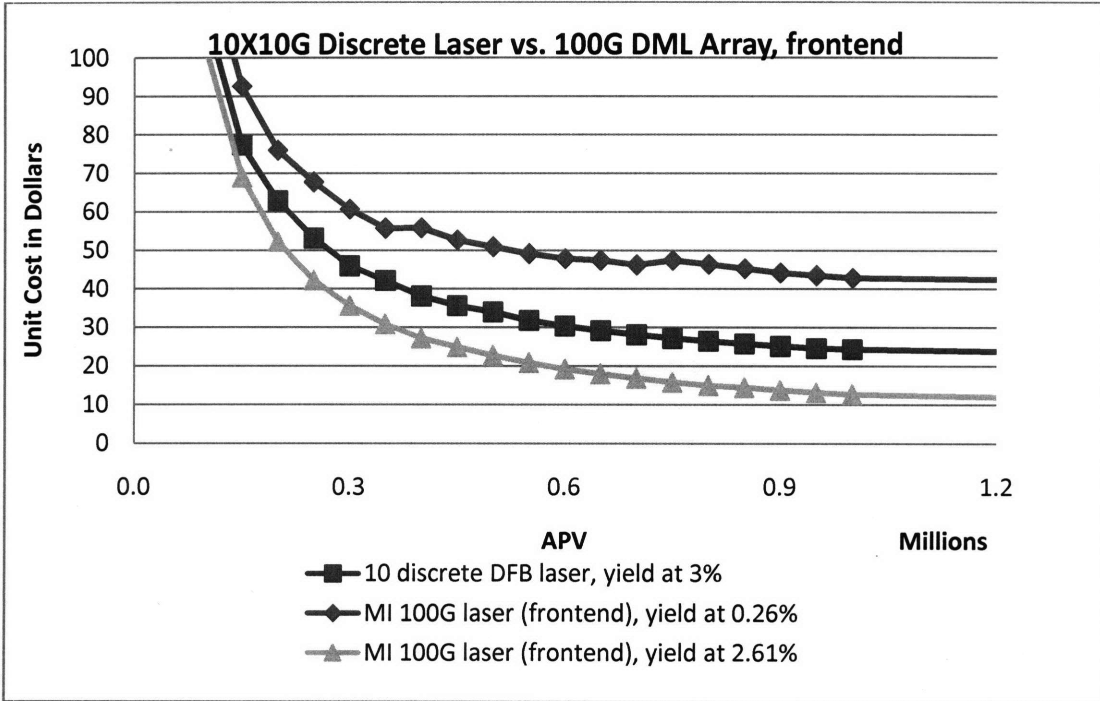


Figure 21. 10 by 10G discrete laser vs. 100G DML array cost comparison, frontend

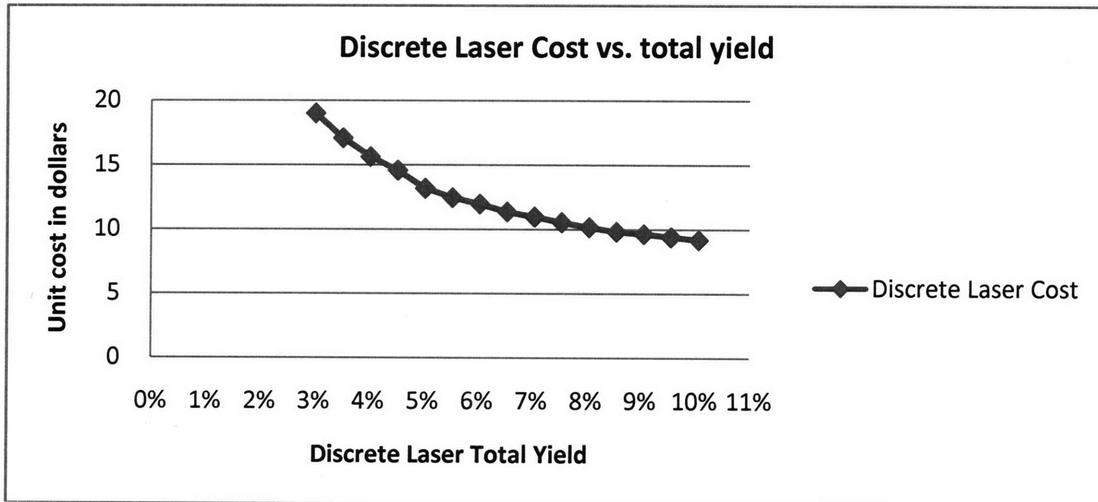


Figure 22. Discrete 10G laser cost vs. total yield, frontend

Further investigation on the sensitivity of the cost estimate to yield at each process step is needed to map out the key drivers of cost. Figure 23 and Figure 24 represent this type of analysis using a uniform 0.1% yield improvement for all processes. The two designs have similar high-impact process steps in slightly different ranking orders. Firms should focus their yield improvement efforts on the highest ranked process steps. For example, a 0.1% yield increase on bar cleave would bring a 0.5% savings in the total frontend cost for the DML Array.

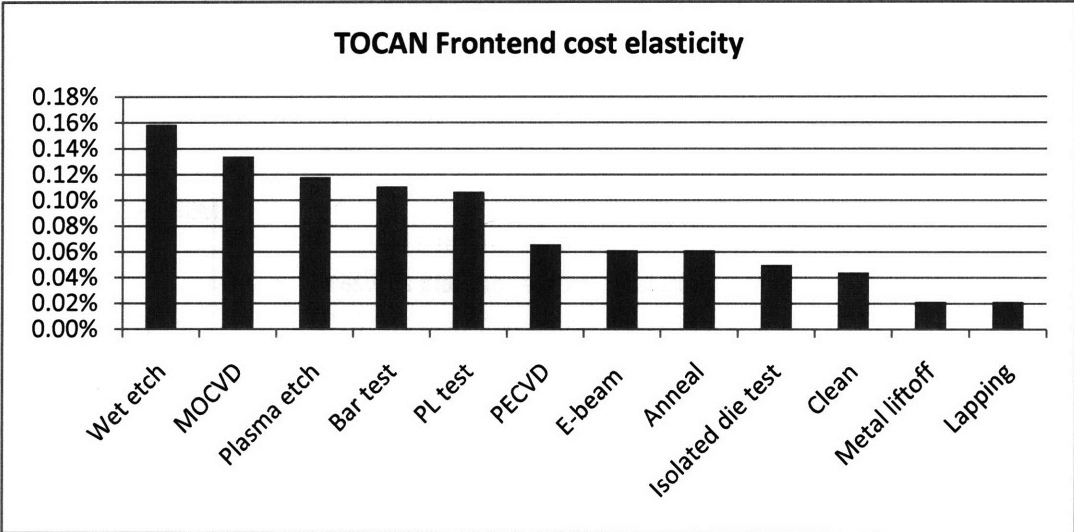


Figure 23. TO-CAN frontend cost elasticity

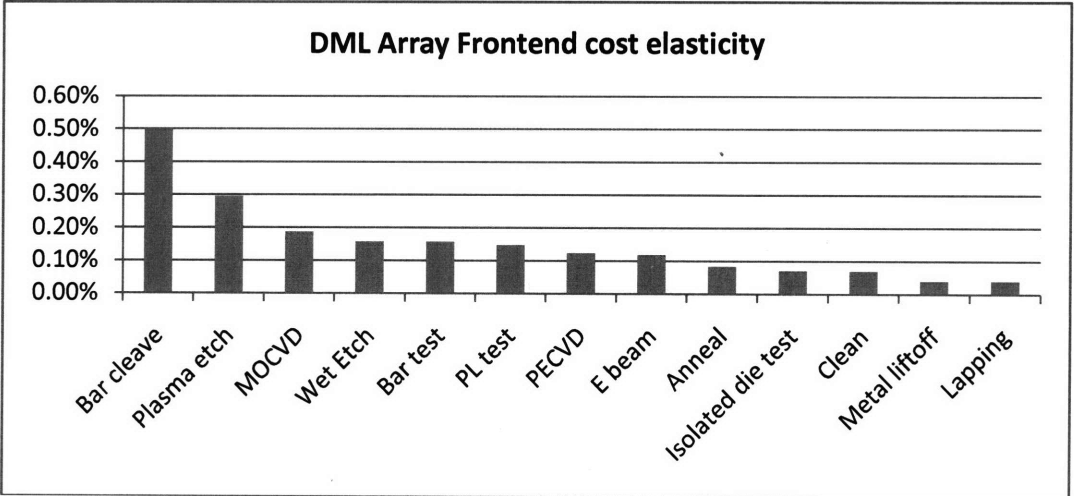


Figure 24. DML array frontend cost elasticity

4.2.1.2 Backend comparison

Backend production processes are defined as the packaging and assembly of chips produced by frontend fabrications into fully functional devices. For the InP designs in this study, backend involves the incorporations of passive photonic functionalities such as the alignment of lenses, prisms, and filters, as well as fitting these into protective housings for the transmitter portion of the transceiver. The receiver portion is not modeled in this study so its cost is included as a constant dollar amount. The backend process flow for Design 1 has 24 steps and Design 2 has 17 steps. Yield assumptions are listed in Table 16.

Table 16. Backend yield assumptions in Design 1 and 2

	Design 1. TO-CAN	Design 1. TOSA	Design 2. DML Array
Backend Yield	95.7%	95.9%	81.7%

4.2.1.2.1 Economy of Scale

The backend production for Designs 1 and 2 demonstrate economies of scale at much lower annual production volumes compared to their frontend, leveling out at around 20,000 APV. This effect is primarily due to lower equipment costs and longer per process cycle time on the backend. The TO-CAN backend stabilizes around $370 per unit, and the DML array backend stabilizes around $55 per unit. This result indicates that the integrated device is always less costly than the aggregate of the discrete devices for the backend. Furthermore, the differential in backend cost is large enough to offset any frontend savings associated with the discrete laser production at reasonable volumes.

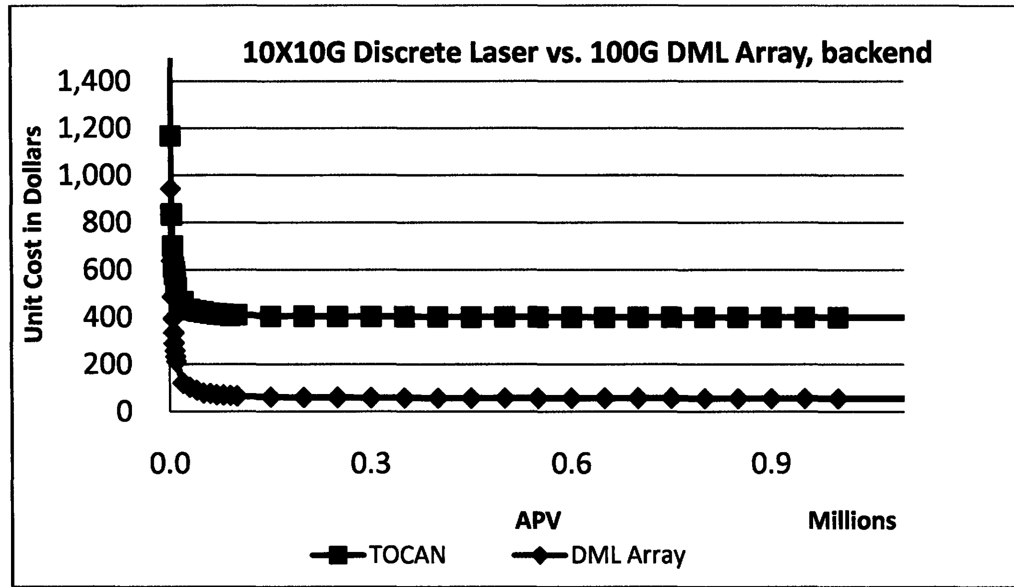


Figure 25. 10 by 10G discrete lasers vs. 100G DML array cost comparison, backend

4.2.1.2.2Top Cost Drivers by Process Type and Cost Element

For the backend, other fixed (tooling), materials, equipment, and labor costs play equally important roles. No single cost element predominates across process types. In general, optical sub-assembly processes are more costly than packaging processes. The backend for the discrete design is much more costly than the backend for integrated design due to complexities associated with individually assembling and aligning ten separate TO-CANs with filters (i.e., weld and filter assembly steps). At 1M APV, total backend allocated investments (equipments and tools) for Design 1 is $348M, and for Design 2 is $48M. Design 1 requires seven times the investment of Design 2!

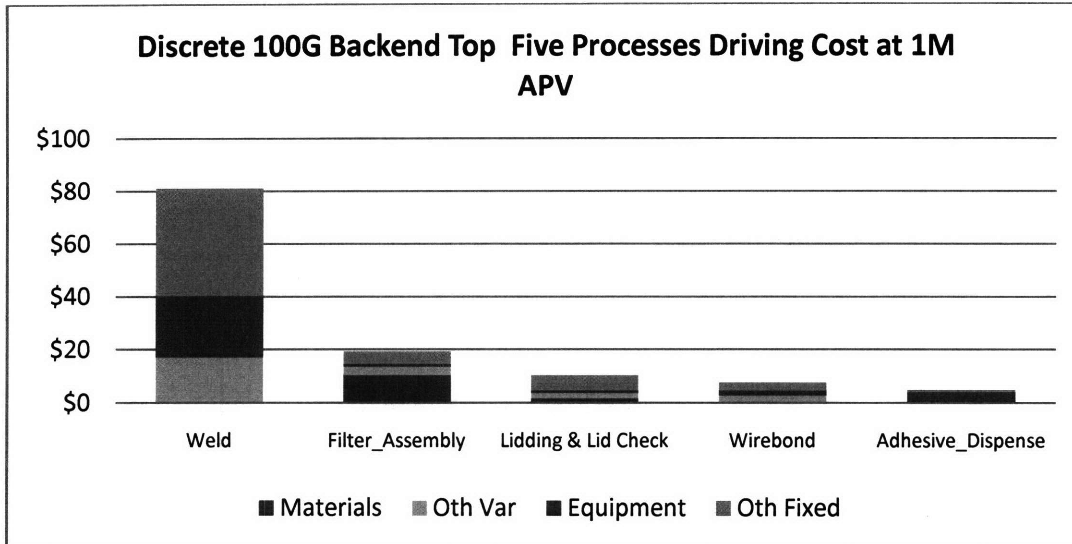


Figure 26. Discrete 100G backend top five processes driving cost at 1 million APV
Total Cost: $128, consisted of one 10G TO-CAN ($27) plus TOSA ($101)

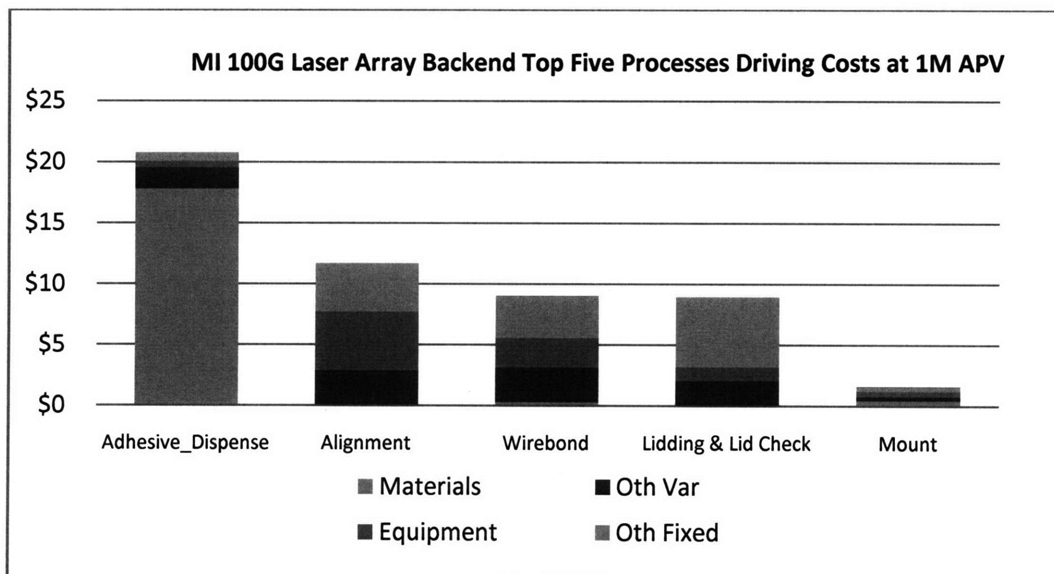


Figure 27. 100G DML array backend top five processes driving costs at 1 million APV
Total Cost: $55

4.2.1.3 Conclusion

For the InP designs, the preceding results suggest that further integration would be expected to provide overall cost savings. This conclusion emerges because the backend cost penalty of discrete lasers overwhelms their frontend cost advantage of the integrated lasers. Design 2's

economic competitiveness is due to cost savings associated with the elimination of costly discrete components and assembly steps.

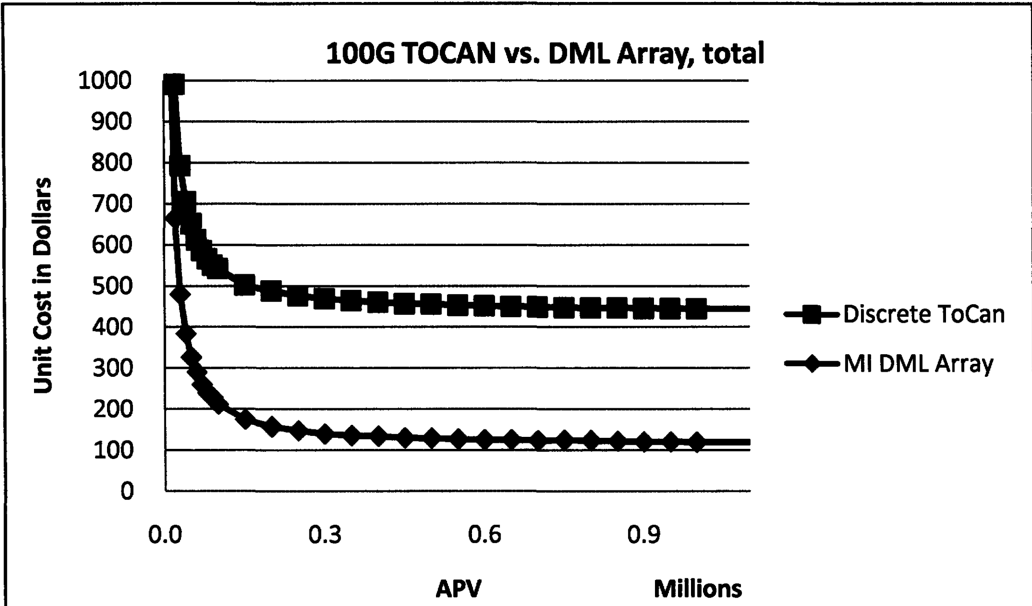


Figure 28. 100G TO-CAN vs. DML array total cost comparison

Figure 28 displays the total cost comparison of Design 1 and 2. At around 0.3 M, both designs have reached economies of scale. The total cost of the TO-CAN design is above $400, and the total cost of the DML Array design is above $100. Integration reduces total cost by a factor of four in this case.

4.2.2 Si Designs

The Si design space includes Design 3 Hybrid and Design 4 Si Two Chip. This section presents a cost analysis and comparison of these two designs.

4.2.2.1 Frontend (Photonic Chips) Comparison

For the Si-based designs in this study, frontend is the production of photonic chips (PC). The frontend process flow for Design 3 has 49 steps and Design 4 has 118 steps. The main difference between these two processes is that the PC of the Si Two Chip design is at a much higher degree of integration compares to the PC of the Hybrid design because it incorporates additional modulation and detection functionalities. Table 17 lists the underlying assumptions in the model:

Table 17. Frontend model assumptions for Design 3 and 4

	Hybrid (8 inch)	Si Two Chip (8 inch)
Die Size	12 mm x 15 mm	14 mm x 15 mm
Die/Wafer	105	90
Material	SOI	
Wafer Cost	$250 (high volume), $300 (low volume)	
Machine Life	5 years	
Frontend Yield	85%	90%

In this comparison, Design 3 and 4 are modeled using data on 8-inch wafer facilities. A 12-inch wafer analysis is conducted in Appendix D. The number of die per wafer is calculated using available wafer areas (taking account of edge exclusion[14]) divided by die size estimated by industry experts. Both designs are built on the same type of raw SOI substrate.

[14] Edge exclusion differs by wafer sizes. 150 mm wafer: 12 mm. 200 mm wafer: 6 mm. 300 mm wafer: 3 mm.

4.2.2.1.1 Economy of Scale

Figure 29 displays the modeled costs for the two types of PCs. Productions of both designs show strong economies of scale up to annual production volumes of approximately 3 million units. At annual volumes above 3 M units, the production costs of the two devices level out, the Hybrid PC at just above $20 per unit, and the Si Two Chip PC at just below $30 per unit. This result indicates that the highly integrated PC is always more costly than the Hybrid PC for the frontend. At high production volumes, the cost advantage for Hybrid becomes very small.

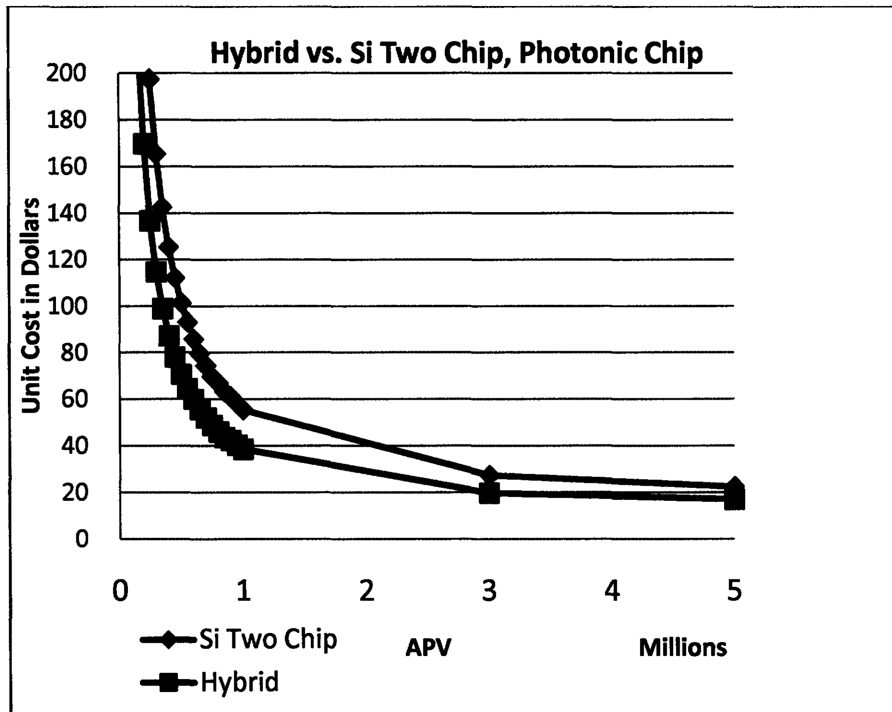


Figure 29. Hybrid vs. Si Two Chip, frontend photonic chip cost comparison

4.2.2.1.2 Top Cost Drivers by Process Types and Cost Element

Figure 30 and Figure 31 present the top seven processes making the greatest contributions to the total cost of the PCs at 10M APV. Both designs share the same top three cost drivers—plasma etch, photolithography, and PECVD—all dominated by equipment and other fixed costs. Testing is also a major cost driver for the Si Two Chip design. At 10M APV, total frontend allocated investment (equipments and tools) for Design 3 is $311 M, and for Design 4 is $414 M. The integrated design requires a higher capital investment due to its lower production yield and additional equipment requirements such as ion implanters and measuring machines.

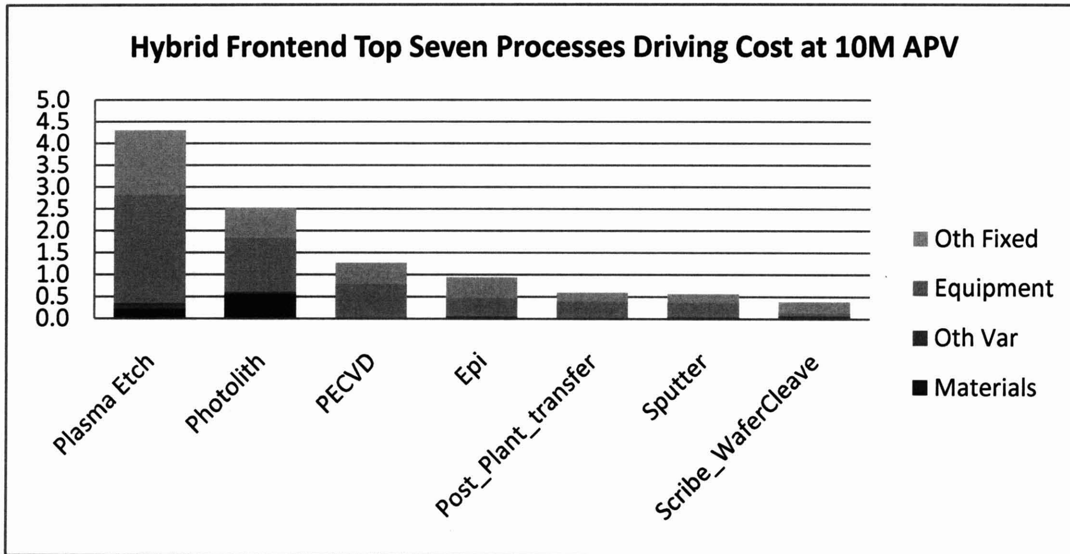


Figure 30. Hybrid frontend top seven processes driving cost at 10 million APV
Total Cost (exclude substrate cost): $11.7

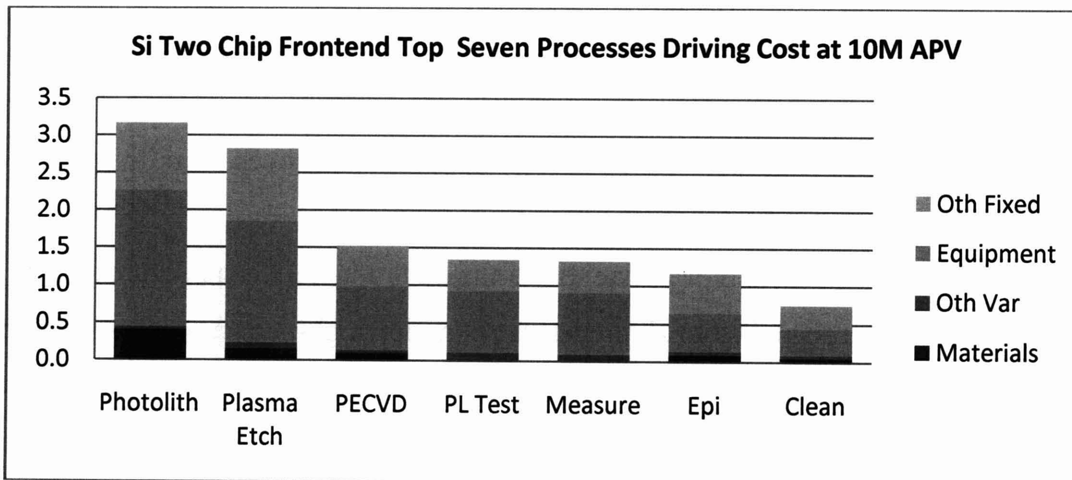


Figure 31. Si Two Chip frontend top seven processes driving cost at 10 million APV
Total Cost (exclude substrate cost): $15.2

4.2.2.1.3 Sensitivity of Cost Estimate to Production and Process Yield

The impact of yield improvements on the frontend production cost of PCs is quantified in Figure 32. This figure displays the costs of a range of yielded PCs for both designs. Hybrid frontend yield ranges from 70% to 95%. Si Two Chip frontend yield ranges from 40% to 95%. After reaching economies of scale, theses result show a lowest cost for Si Two Chip at 95% total yield, and a highest cost for Si Two Chip at 40% total yield out of the six variations. It is interesting to note that a 5% to 10% total yield increase does not lead to significant cost savings

for both designs. This result is consistent with the last segment of the technology S-curve in which continuous improvement of mature technologies gives diminishing returns.

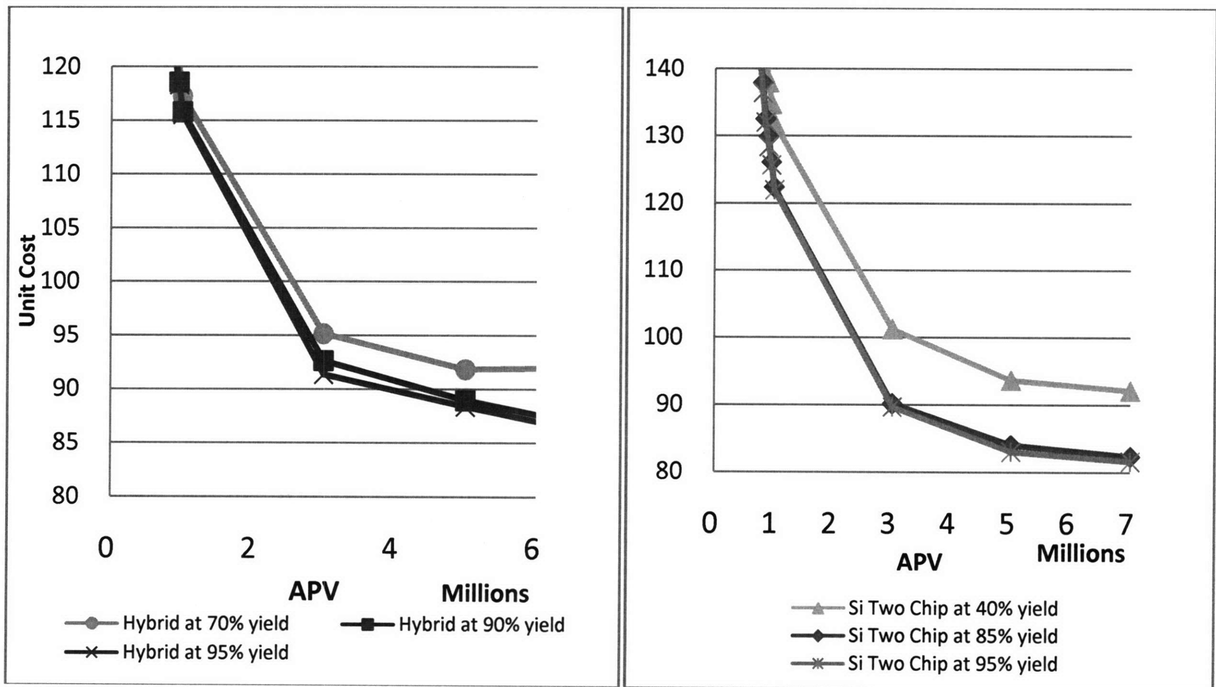


Figure 32. a) Hybrid total cost, sensitivity of APV to total yield, and b) Si Two Chip total cost, sensitivity of APV to total yield

Sensitivity of cost estimates to process yields are shown in Figure 33 and Figure 34. A uniform 0.1% yield change is applied to all processes. Common high-impact process steps are plasma etch, photolithography, and PECVD. Thermal, wet etch, and sputter are important steps to Design 3, while testing steps and wet processes are important steps to Design 4. These results show a different ranking order than Figure 30 and Figure 31. The top cost drivers shown in the previous section emphasize the aggregate cost contributions by each process, where the results in this section emphasize the individual process's yield impact on the total cost.

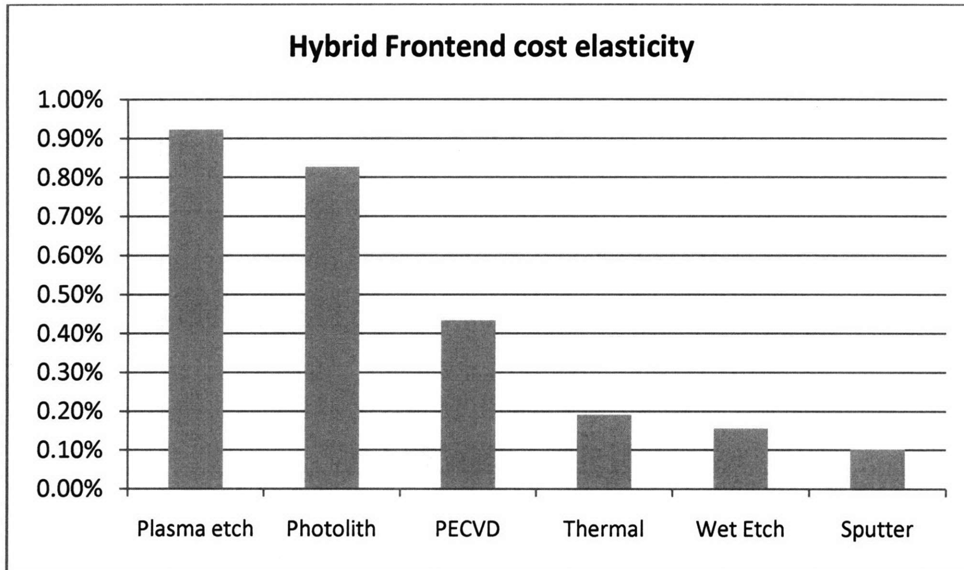


Figure 33. Hybrid frontend cost elasticity

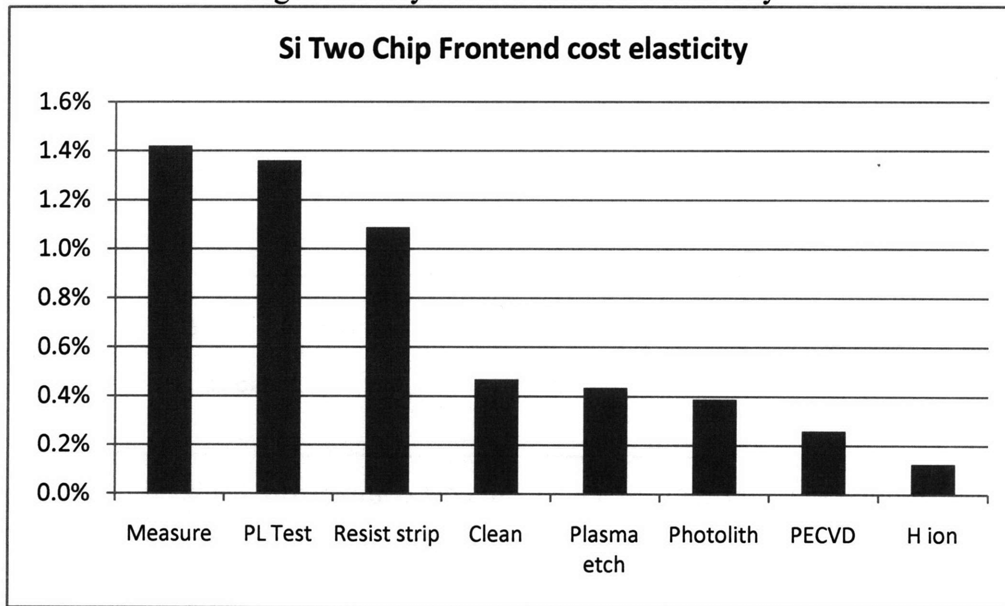


Figure 34. Si Two Chip frontend cost elasticity

4.2.2.2 Backend Comparison

The Hybrid and Si Two Chip designs use very similar backend processes. Backend involves laser attachment and lens alignment, as well as fitting into protective housings. The main difference between the two designs is that the Hybrid design requires an additional photo-detector attachment step. Backend process flow for Design 3 has 17 steps and Design 4 has 16 steps. The backend yield for both designs is modeled at 87.4%.

4.2.2.2.1 Economy of Scale

The backend productions for Design 3 and 4 demonstrate much faster economies of scale compared to their frontend, leveling out around 0.15 M APV (Figure 35). This is primarily due to lower equipment costs and longer per process cycle time on the backend. The Hybrid backend stabilizes around $35 per unit, and the Si Two Chip backend stabilizes around $27 per unit. This result indicates that the Hybrid is always more costly than the Si Two Chip for backend assembly.

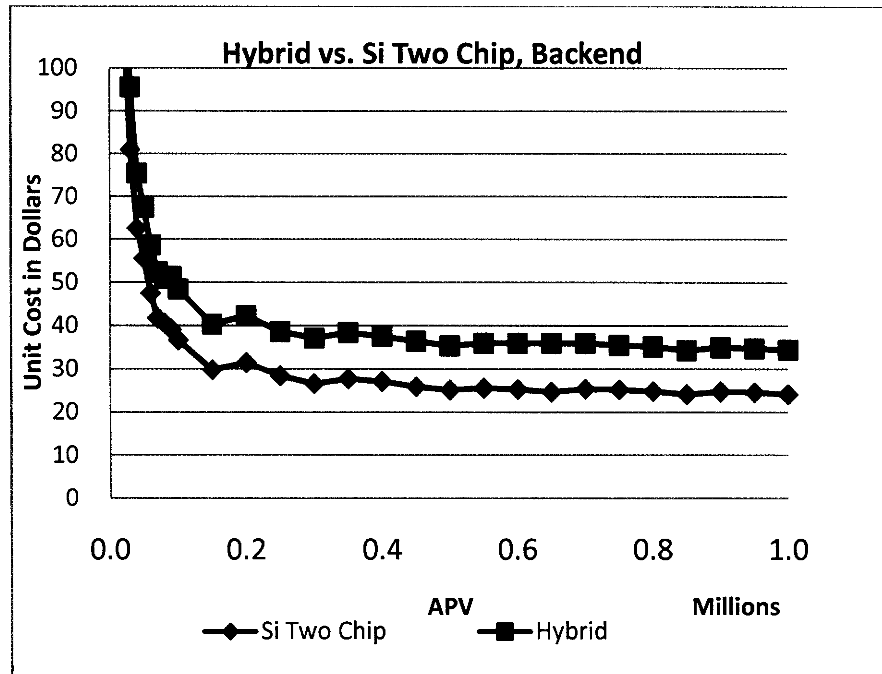


Figure 35. Hybrid vs. Si Two Chip cost comparison, backend

4.2.2.2.2 Top Cost Drivers by Process Types and Cost Element

Figure 36 and Figure 37 present the top five processes making the greatest contribution to the backend cost at 10M APV. Both designs share the same top cost drivers, chip bond, wirebond, die polish, assembly test, and alignment. The only difference is that the ordering of chip bond and wirebond is switched due to the additional photo-detector attachment step (modeled as one type of chip bond) in the Hybrid design. At 10M APV, total backend investment (equipments and tools) for Design 3 is $372 M, and for Design 4 is $311 M. The Hybrid design requires a slightly higher capital investment due to the additional time and capital required for the photo-detector attachment step.

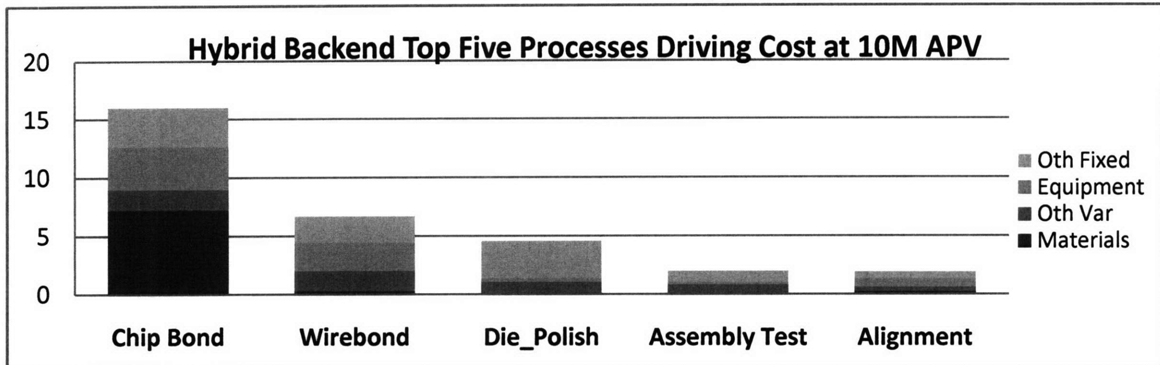


Figure 36. Hybrid backend top five processes driving cost at 10 million APV
Total Cost: $33.5

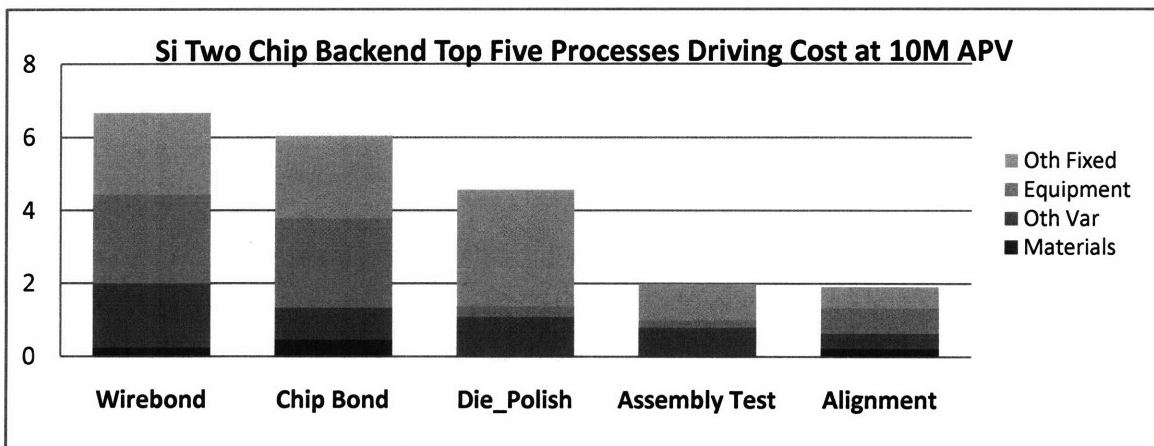


Figure 37. Si Two Chip backend top five processes driving cost at 10 million APV
Total Cost: $23.5

4.2.2.3 Conclusion

In the Si design space, further integration provides overall cost savings at a volume greater than 3 M units per year. This conclusion is reached based on the fact that frontend fabrication of the photonic chip in the Hybrid design is more cost competitive than the monolithically integrated photonic chip in the Si Two Chip design, while the backend cost of the Hybrid design is slightly higher than the Si Two Chip design. Design 4's economic competitiveness at high volume is due to cost savings associated with the elimination of discrete components and assembly steps (Figure 38).

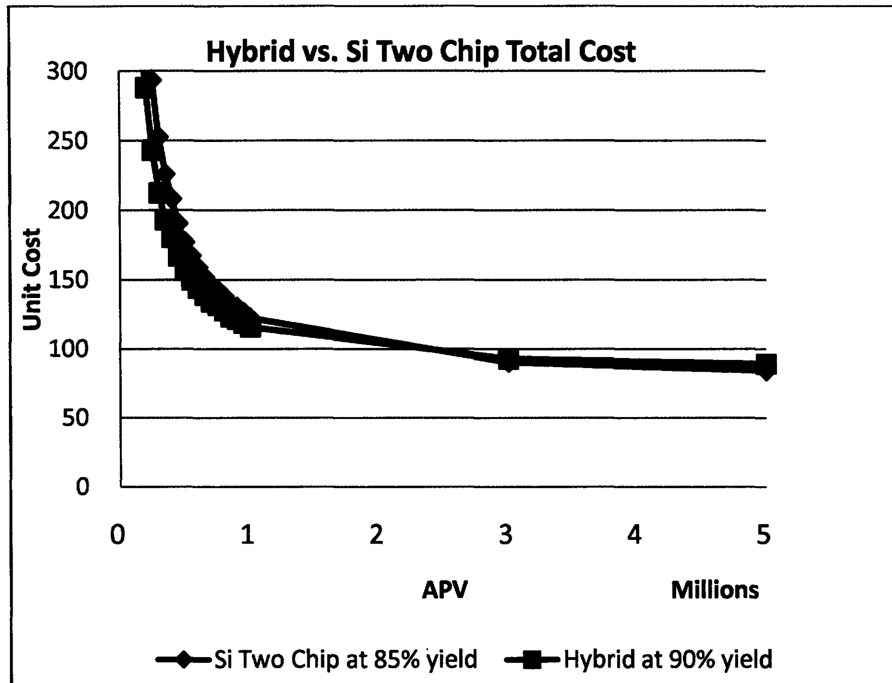


Figure 38. Hybrid vs. Si Two Chip total cost comparison

4.2.2.4 Cost Estimation on Si One Chip Design

The cost for the Si One Chip Design introduced in Section 4.1.4.1 can be roughly estimated by common rules used in the semiconductor industry. As die size increases with additional optics device integration, two effects can occur: 1) dies per wafer are reduced, and 2) the larger die will experience a lower yield on a percentage die per wafer basis. The rule of thumb for a 10mm square die size is Δ% die Yield = 0.22 x [% increase in die size[15]]. Assume in the One Chip solution the integrated photonic components will make up two-third of the new die[16]. For an integrated die with two-third photonic real estate,

$DPW_{integrated} = \frac{10 \times 10}{17 \times 17} \times (1-0.22*70\%) = 29.2\%$ of $DPW_{CMOS\ only}$. DPW is the abbreviation for device per wafer (Haubensak 2007).

The yielded DPW for a CMOS chip on a 200mm wafer is approximately 207. The new DPW for the One Chip design is 29.2% x 207 = 60. This number accounts for the yield hit associated with the die size increase alone. To account for additional yield hit with photonic processing, the DPW of 60 is used as an input to the FOP-PBCM. The die cost for the Si One

[15] Die size is defined as the length of one side of a square die

[16] CMOS die: 10mm x 10mm, photonic die: 14mm x 15mm, integrated die: 17mm x 17mm

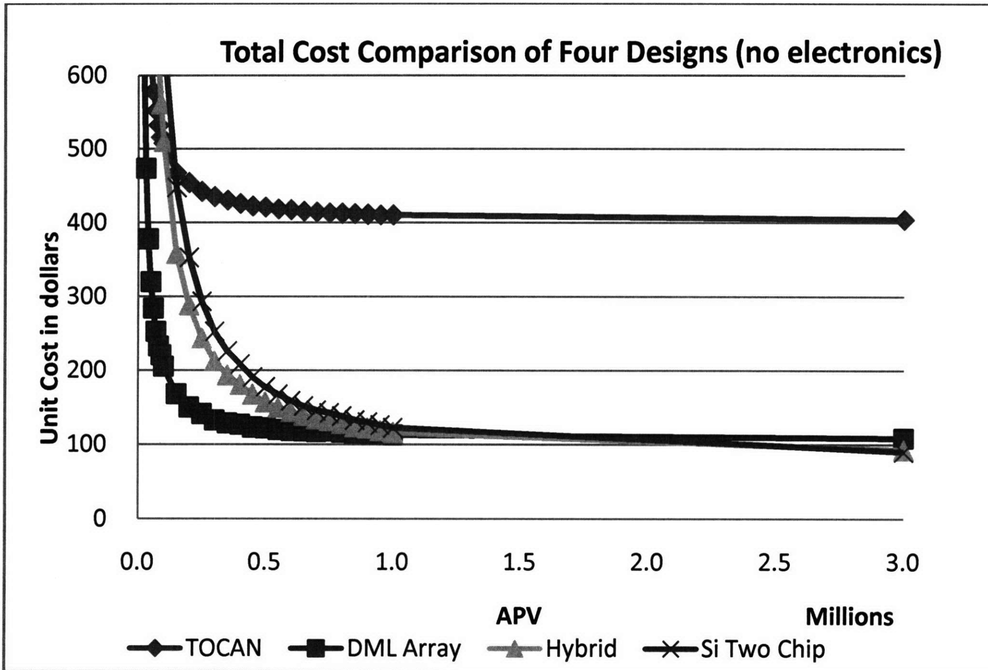
Chip design is $27 at 10M APV from the output of the model, which is roughly 1.35 times the cost of the Si Two Chip design. This number can be seen as a lower bond cost estimate because systemic yield hit and additional complications with electronic-photonic integration are not accounted in this analysis. For instance, "the impact on yield will depend on the temperature used and the transistor capability requirement. If high thermal treatments are used to minimize optics performance, a worse case 25-50% impact could occur from the transistor degradation alone" (Haubensak 2007).

4.3 Cost Comparison of Material Platforms

A cost comparison of all four designs provides some insights for material platform selection. After all designs have reached economies of scale, further integration provides overall cost advantages. At a volume greater than 2.5 to 3 million units per year, the following cost relationship applies: TO-CAN > DML Array > Hybrid > Si Two Chip. The ">" sign indicates "more costly" in this equation. However, economic competitiveness is highly dependent on volume expectations. The TO-CAN design is never cost competitive. Multiple cross-over points exist among DML Array, Hybrid, and Si Two Chip shown in Figure 39. Figure 39 b) shows a delta cost graph using the cost of Si Two Chip as the x-axis. The lowest curve is the most cost competitive solution:

Winner at	Low (less than 1.1 M):	DML Array
	Medium (1.1 M to 2.5 M - 3 M):	Hybrid
	High (greater than 3 M):	Si Two Chip

At low volume, cost estimates would indicate that the InP material platform is preferred, while at high volume the Si material platform is preferred. Backend plays a major role in determining cost competitiveness (Figure 40). The TO-CAN design is never cost competitive among the four designs due to its large gap in backend cost.



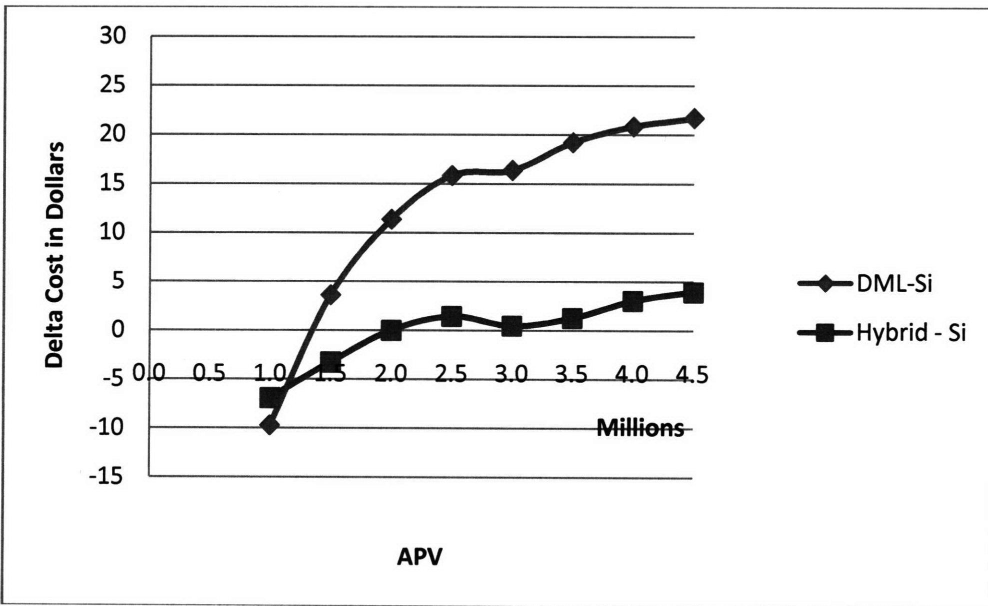


Figure 39. a) Total cost comparison of four designs, b) delta cost using Si Two Chip as the baseline at high volume

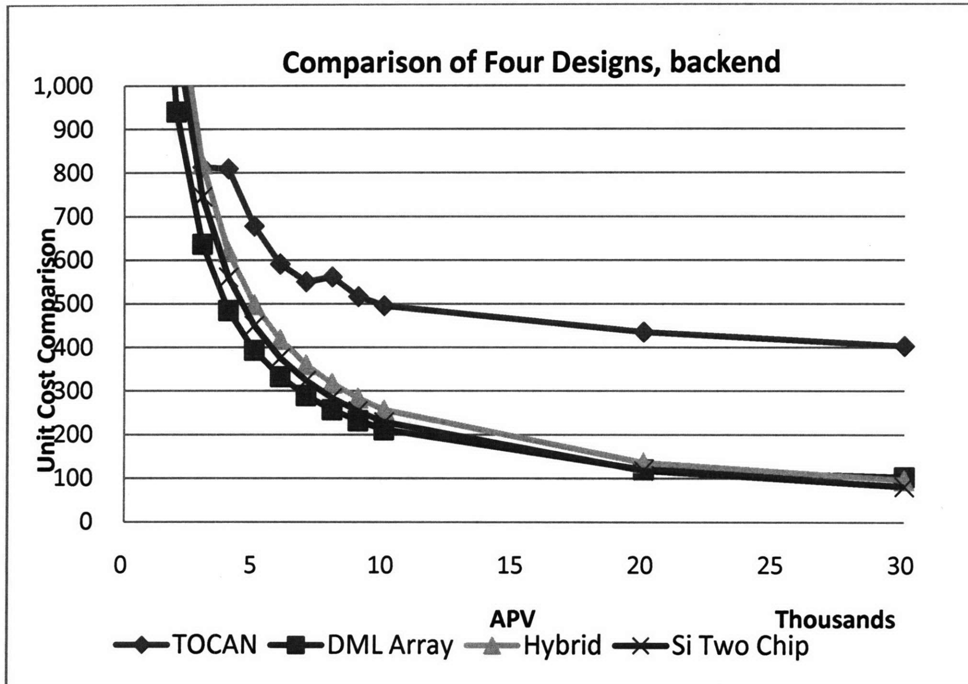


Figure 40. Cost comparison of four designs, backend

4.4 Analysis on Three Factor Comparison

As discussed in Section 3.5.2, this research gathered market size projection data to provide context to the previous cost results through the collective roadmapping effort at the MIT MPC. The goal for this research is to produce a three-factor analysis, mapping out production cost for each transceiver design and volume combination, and then comparing these costs with price expectations to determine the viability of such transceiver markets in the datacom and computing industries (Figure 41).

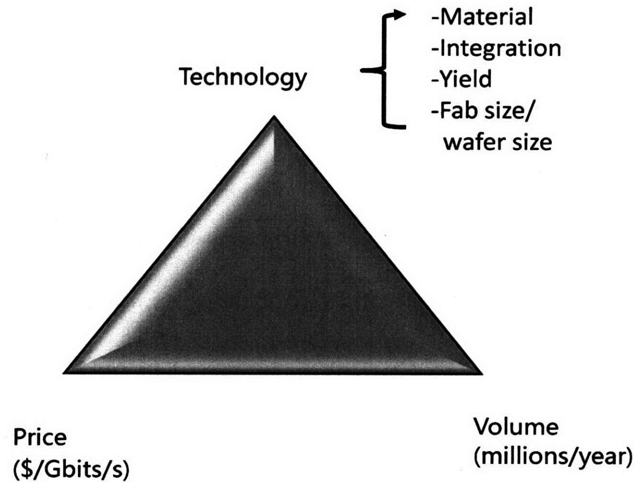


Figure 41. Three Factor Analysis

Initial results are averaged in Table 18 and Table 19 from the first round of cross-market interviews. Six companies participated in this process for the datacom and computing markets spanning the entire value chain (chip, component, and system manufacturers). In Table 18, "link" is defined as a dedicated transmit or receive signal on a single fiber, and "system" is defined as a full rack containing 40 blade servers or cards in a typical data center. In Table 19, optical server volume is obtained by multiplying server volume and expected percentage optically connected systems. These results are then divided by 40 to obtain the optical system volume. In Table 20, 100G transceiver volume and the trigger price for 100G are calculated from the aggregate bandwidth and $/Gbit/s in the above tables. The costs for the TO-CAN, DML, Hybrid, and Si designs are outputs of the FOP-PBCM model at each 100G transceiver volume, assuming baseline yields for all designs and 8 inch fab lines for Design 3 and 4.

Table 21 shows the difference between the trigger price and modeled costs at the expected market size for each of the four years analyzed.

Table 18. Cross market survey results, datacom

Year	Data rate per link (Gbits)	# Link/system	Data rate per system (Gbits)	Data rate per system (Tbits)	$/Gbit/s
2007	11	158	1,500	1.5	$4.40
2010	36	634	12,908	12.9	$2.00
2013	72	1,640	52,333	52.3	$1.15
2016	122	4,348	247,500	247.5	$0.74

Table 19. Server volume projection
(IT Hardware Research 2007)

Year	Server volume (k)	% Optical	Optical server volume	Optical system volume	Aggregate BW (Tbits)
2007	7,100	1%	71,000	1,775	2,663
2010	9,450	10%	945,010	23,625	304,963
2013	12,578	20%	2,515,617	62,890	3,291,265
2016	16,741	35%	5,859,500	146,488	36,255,656

Table 20. Transceiver volume, trigger price, and PBCM cost results across four designs

Year	100Gbits transceiver volume	Trigger price for 100Gbits	TO-CAN	DML	Hybrid	Si
2007	26,625	$440	$758	$473	$1,557	$2,040
2010	3,049,626	$200	$404	$109	$86	$90
2013	32,912,651	$115	$402	$106	$84	$77
2016	362,556,563	$74	$402	$106	$84	$77

Table 21. PBCM cost results across four designs minus trigger price

Year	TO-CAN - TP	DML - TP	Hybrid - TP	Si - TP
2007	$318	$33	$1,117	$1,600
2010	$204	-$91	-$114	-$110
2013	$287	-$9	-$31	-$38
2016	$328	$32	$10	$3

Figure 42 plots the results in Table 21. In all years, the TO-CAN design is not viable for the datacom market. Designs 2, 3, and 4 all met the trigger price expectation in the years 2010 and 2013. However, taking into account that the PBCM cost numbers exclude electronic costs, the actual total costs for such devices are likely to be much higher (30% to 60% higher). In this case, the total cost of the DML design in 2013 is likely to be above the trigger price. As time goes on, trigger price decreases at a faster rate than economies of scale in production. Designs below the trigger price in 2013 are no longer viable in 2016.

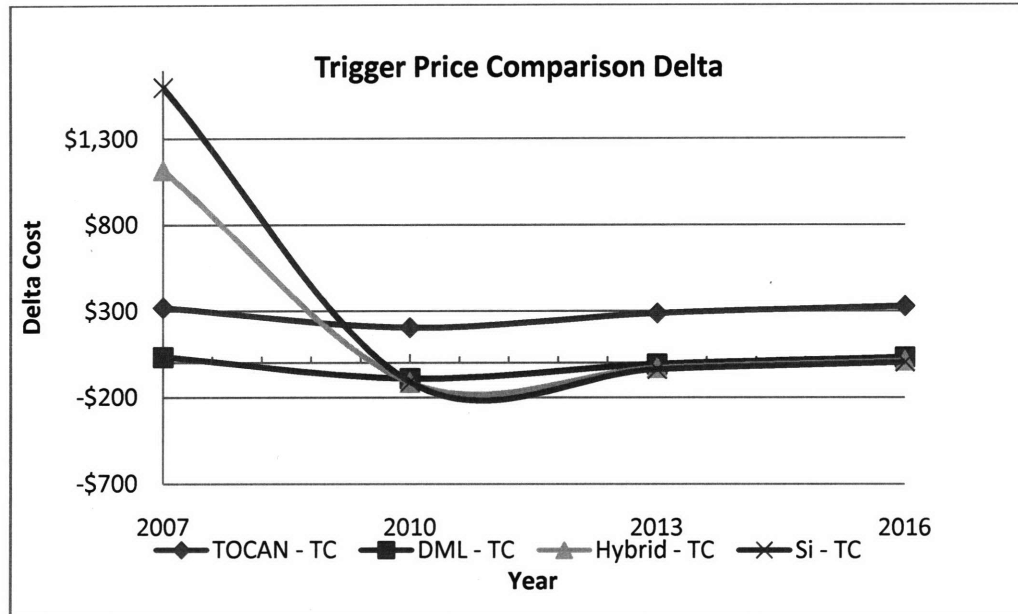


Figure 42. Trigger price comparison delta graph, datacom

A similar analysis is conducted for the computing market (Table 22). The meaning of Personal Computing (PC) here includes desktops, laptops, and servers. Interconnects for the computing market are mainly used for chip-to-chip communication. This analysis only compares the frontend cost of the Si Two Chip design with the trigger price due to the architecture of chip-to-chip communication. Design 1 to 3 are more discrete designs that are not viable solutions on the inter- and intra-chip level. Packaging cost is also excluded because the type of packaging, if any, used for chip-to-chip communication is significantly different from the backend flow modeled in PBCM, and it is likely to be at a much lower price point.

Table 22. Cross market survey results, computing (chip-to-chip).
(IT Hardware Research 2007)

Year	2007	2010	2013	2016
The number of PCs grows by 10% each year. Optical penetration: 1% in 2007, 15%-20% in 2013, 30%-40% in 2015 and later.				
PC volume (k)	260,862	347,207	462,133	615,099
%Optical	1%	10%	20%	35%
Optical PC volume	2,608,620	34,720,732	92,426,589	215,284,633
Data rate (Gbits)	15	40	120	240
Aggregate BW (Gbits)	39,129,300	1,388,829,288	11,091,190,694	51,668,311,848
100Gbits transceiver volumes	391,293	13,888,293	110,911,907	516,683,118
Trigger Price for 100Gbits	$40	$30	$20	$10
Si Two Chip Frontend Cost	$125	$18	$16	$15
Si Two Chip Cost – TP	$85	$-12	$-4	$5

Figure 43 displays the comparison result for the computing market. It indicates a window of opportunity for Si photonics to enter this space between 2010 and 2013. However, Si photonic technology must achieve a relatively mature yield (85% yield and above) by that time for this conclusion to be valid. In addition to technical challenges, a host of organizational and market challenges still remains in the commercialization of Si photonics. To enable electronic-photonic integration on the chip level, the semiconductor industry must be committed to Si photonics. Challenges and issues on industry coordination and standards across markets will be addressed in the concluding chapter of this thesis.

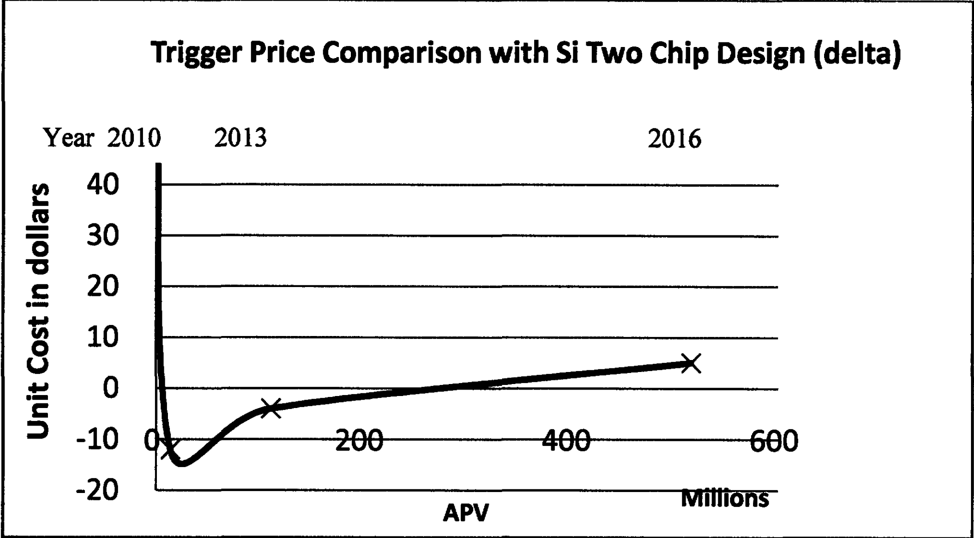


Figure 43. Trigger price comparison with Si Two Chip (delta), computing

5 Conclusion and Future Work

In recent years, consumers demand ever increasing processing power, storage capacity, and I/O capacity in personal computing, data network, and display interface. At the same time, the existing copper interconnect infrastructure in the datacom and computing industries is likely to reach performance limitations and create a severe communication bottleneck problem in the near future. Optical interconnects may be the ideal substitute for copper with the potential to create a new communication platform using planar waveguides on the micro- and nano-scale. For this transition to take place, significant technological, market, and organizational challenges must be addressed by the converging optoelectronic and semiconductor IC industries.

The optoelectronic industry is predominately built on the III-V material platform. However, research advancements of this decade give new promise to Si as a viable photonics platform. This is significant due to its potential to enable large scale electronic-photonic integration, and the reuse of the CMOS manufacturing infrastructure. Using emerging 100G Ethernet LAN transceivers as a case study, this research applies the process-based cost modeling method to characterize the production cost of alternative designs on both of the InP and Si material platforms. In each material platform, one discrete and one integrated designs are modeled using data collected from firms across the industry supply chain. The following conclusions answer the three research questions proposed in this thesis by discussing the potential for Si photonics to displace III-V technology and invade multiple technically converging markets, and how standards and industry coordination can enable technology adoptions in the optoelectronic industry. This chapter also discusses major thesis contributions and prospects for future research.

5.1 Research Conclusions

Technology question: what is the most cost competitive architecture and material solution in manufacturing 100G Ethernet LAN transceivers?

Four designs were examined in this research: TO-CAN and DML Array on the InP material platform, and Hybrid and Si Two Chip on the Si material platform. Within each material platform, integration is cost competitive. This economic competitiveness is due to cost savings associated with the elimination of discrete components and assembly steps. On the frontend, the discrete designs have slight cost advantages compared to the integrated designs due to their higher yields. However, it is not necessary for the integrated designs to reach yield parity to match and drop

below the frontend cost of the discrete designs due to their backend cost savings. A total cost comparison across material platforms indicates at low volume (less than 1.1 million annual units) the InP material platform is preferred, while at high volume (greater than 3 million annual units) the Si material platform is preferred. The Si based transceiver requires volume manufacturing to recover heavy frontend capital investments (i.e. equipment and lithography mask tools).

For Si photonics to displace III-V photonics, the cost of achieving a viable technology is not the biggest barrier: Si photonics demonstrated the ability to integrate detectors at low cost, material and process compatibilities with CMOS (Jalali and Fathpour 2006), and the light source may become an off-chip power source in future interconnect configurations. There are still concerns with a Si photonic device's heat compatibility with hot VLSI processing that are not addressed by this thesis (Jalali and Fathpour 2006). Assuming that is possible, the bigger barriers for Si photonic technology gaining commercial acceptance are market and organizational issues. This research indicates that the Si based transceiver is cost competitive at an annual volume greater than 3 million units per year per fab. The MIT CTR's Si TWG suggests that a market of 50 million is required to support 3 million annual units per fab. The next section will address this market challenge.

Market question: what is the structure of the three factors: cost of achieving the technology, volume expectation, and required substitution price that will characterize an advantageous state for Si photonics to enter the Datacom and Computing markets?

The three factor analysis as the result of the MIT CTR cross market survey indicates that annual production volumes must be in the tens of millions unit range to provide the minimum economies of scale necessary for the designs to meet the trigger price expectation. Volume in the tens of millions range requires a paradigm shift from the typical telecom market mindset held by the optoelectronic industry, and a move toward the mindset of the semiconductor IC industry. To provide some context to these ranges, in the server markets, a server is sold every 12 seconds on a 24/7 scale, which translates to a volume of 2.6 million annual units (Si Technology Working Group 2008). Depending on the I/O capacity and percentage of optical interconnects used in each server, this server volume implies a much larger interconnect volume that can approach the 50 million required annual volume for Si photonics to present a cost advantage. Therefore, although market volume is a barrier for the adoption of Si photonic interconnects, the biggest

barriers may lay in the organizational structures of these converging industries. In addition, trigger prices are decreasing at a rate faster than cost reduction rate through economies of scale. No design examined in this thesis can achieve a cost lower than the price expectation in year 2016 given current processing conditions. This conclusion implies that firms must achieve higher yields on their proposed interconnect technologies, and/or continue to innovate in product design and manufacturing.

Organization question: what roles could industry standards and coordination (roadmapping) play as enablers of emerging optoelectronic interconnect technology across markets?

Results of the cost modeling study and observations from the MIT CTR roadmapping activities indicate that standards and a set of common language are essential to enable converging technology markets. First, product and process standardization are required to enable a Si photonic fab with minimum economies of scale of 3 million annual units and a total addressable market of 50 million annual units. This is very different from the Telecom mindset of focusing on innovation, volume in the tens of thousands range, product design proliferation, and labor-intensive fabrication lines, which can only survive so long as high prices are paid for telecom components (Ragona 2001).

Second, the first move toward industry coordination is effective communication using a set of common language. Participation in the cross market survey as a part of the roadmapping activities shows that significant communication barriers exist between companies occupying different positions on the supply chain. The system, component, and IC engineers don't use the same vocabulary for the same concepts. Furthermore, even the "same" words often have different meanings. For example, the work "link" can mean one unidirectional fiber, one bidirectional fiber, one fiber with multiple channels, two fibers, ribbon fibers, etc. In addition, a successful cross market questionnaire requires participants to dedicate sufficient time and energy as well as employ system thinking. For example, the questionnaire asks four interrelated questions: the data rate per link, number of links per system, channels per link, and number of links per transceiver. An interviewee answering these questions must realize the interdependencies between these elements (some answers may stay the same over the years while others increase). Otherwise, the resulting market projections can be grossly exaggerated.

Third, significant financial barriers exist for the current Si photonic firms. There is a saying that the way to make a small fortune is to start with a big fortune. Unfortunately, the most innovative firms in the photonic industry are currently small and lacking the necessary financial resources to ramp up Si photonic production to high volume even if the market exists. The development and ramp up cost itself can bankrupt a small firm. In the world of semiconductor IC, established PC and servers makers such as HP and IBM depend on two or three large microprocessor and memory chip companies to supply their chips. Photonic technology has to capture the imagination of these dominate IC manufacturers and overcome great resistance to change. Therefore, the current photonic roadmapping efforts are staying away from making broad conjectures on a timeline for monolithic integration of photonics with electronic chips until high level management in these established companies are ready to hear such "electro-political" claims. One strategy of the MIT CTR is to make a strong case for Si photonic technology outside of the computing space and entice the audience of leading IC companies once the technology becomes attractive (Si Technology Working Group 2008).

5.2 Thesis Contribution and Future Research

This thesis aims to provide insights for strategic decision making in optoelectronic firms, and identify obstacles for emerging technology to gain commercial acceptance by identifying potential disruptive innovations, exploring alternative technological paths, and investigating ways to establish industry coordination.

The cost modeling comparison across four 100GE transceiver designs on two material platforms has produced useful insights. To make this research more comprehensive, two areas deserve further in-depth study. First, economic modeling of a monolithically integrated InP photonic chip can provide a more robust analysis to answer the question of cost competitiveness of the InP and Si material platforms. A monolithically integrated InP transceiver has advantages in integrating the light source with rest of the photonic circuits, which may significantly reduce backend packing and assembly cost. However, this design still suffers most of the downside of the discrete III-V photonic industry, such as small wafer size and manual production lines. A second area of future research concentrates on the cost savings from energy usage. One enabling attribute of Si photonics is low power usage. The MIT CTR Si TWG speculated the power consumptions could be 100 times better than the best copper technology (Si Technology

Working Group 2008). Since data center energy usage accounts for 2% of the total U.S. energy consumptions, legislative energy saving policies may accelerate photonic interconnects' commercial adoptions. Studies in this area could provide important insights at the intersection of technology and policy research.

Reference

Bautista, J., M. Morse, et al. (2005). Microphotonics: Hardware for the Information Age—Silicon Microphotonics. Cambridge, MIT Microphotonics Center Industry Consortium.

Beals, M. (2007). Personal Interview. S. Liu. Cambridge.

Bruce, E. J. and C. H. Fine (2007). Technology Roadmapping: A Tutorial. Cambridge, Massachusetts Institute of Technology.

Christensen, C. M. (1997). The innovator's dilemma : when new technologies cause great firms to fail. Boston, Mass., Harvard Business School Press.

Clairardin, X. (2007). Roadmap to 100 GbE, a CWDM Solution. IEEE 802.3 HSSG, Monterey, CA.

Clayton, R. and T. Dudley (2005). Microphotonics: Hardware for the Information Age—Integration in III-V Materials, MIT Microphotonics Center Industry Consortium.

Cole, C. (2007). Broad Market Potential of 100GE Transceivers. IEEE HSSG.

Cole, C. (2007). Technical Feasibility of SMF & MMF 100GE Transceivers. IEEE 802.3 HSSG, Monterey, CA.

Cole, C. and B. Huebner (2007). Common LAN WDM Grid Proposal for all 100GE SMF Reaches. IEEE 802.3 Higher Speed Study Group.

E.A. (Gene) Fitzgerald, M. J. M., C.L.Dohrman, K. Chilukuri (2007). Progress in Monolithic III-V/Si and towards processing III-V Devices in Silicon Manufacturing. MIT Microphotonic Consortium, Communication Technology Roadmap Fall Meeting, Cambridge, MA.

England, J. S. and R. W. England (1998). The Reliability Challenge: New Materials in the New Millennium, Moore's Law Derives a Discontinuity. Reliability Physics Symposium Proceedings 36th Annual 1998 IEEE International Reno, NV, USA, Texas Instruments Incorporated.

Field, F., R. Kirchain, et al. (2007). "Process cost modeling: Strategic engineering and economic evaluation of materials technologies." Jom **59**(10): 21-32.

Field, F. and L. Ng (1989). "Materials for Printed Circuit Board: Past Usage and Future Prospects." Mater. Soc. **13**(3): 301-318.

Ford, M. and Cisco Systems Inc. (1997). Internetworking technologies handbook. Indianapolis, IN, Cisco Press : New Riders Publishing.

Fuchs, E. and R. Kirchain (2005). Understanding the Economics of Integration: Process-based Cost Modeling of Optoelectronics Production. MIT Microphotonics Center Industry Consortium.

Fuchs, E. and R. Kirchain (2006). Consolidate? Integrate? Go East? Or Get Out? Mapping the Cost Drivers in Optoelectronics Production. Cambridge, MIT.

Fuchs, E. R. H., E. J. Bruce, et al. (2006). "Process-based cost Modeling of photonics manufacture: The cost competitiveness of monolithic integration of a 1550-nm DFB laser and an electroabsorptive modulator on an InP platform." Journal of Lightwave Technology **24**(8): 3175-3186.

Hartman, R. (2007). Photonic Integration Circuit (PIC) Alternatives for 100GE. IEEE 802.3 HSSG, Monterey, CA.

Haubensak, F. (2007). CMOS photonic integration yield impact assessment. Cambridge.

HSSG, I. (2007). Agenda and General Information. IEEE 802.3 Higher Speed Study Group Plenary Meeting, Atlanta GA.

ICEC (1997). Cost Effective IC Manufacturing 1998-1999: Chapter Seven "Changing Wafer Size and the Move to 300mm", Integrated Circuit Engineering Corporation.

Intel. (2008). "International Technology Roadmap for Semiconductors." 2007, from http://www.intel.com/technology/silicon/itroadmap.htm.

Intel. (2008). "Standards & Initiatives." 2007, from http://www.intel.com/standards/index.htm?iid=homepage+hdr_nav2_standards.

IT Hardware Research (2007). UBS Global PC Unit Growth Forecast, UBS.

Jaeger, J. and D. Perkins (2007). 100G Ethernet, Technical Feasibility & Reliability Support for WDM SMF PHY Approaches. IEEE 802.3 HSSG, Monterey, CA.

Jalali, B. and S. Fathpour (2006). "Silicon photonics." Journal of Lightwave Technology **24**(12): 4600-4615.

Kanellos, M. (2002, March 5). "Intel to Make Optical Chips for Others." Retrieved October 1, 2007, from http://www.news.com/Intel-to-make-optical-chips-for-others/2100-1033_3-851459.html.

Khodja, S. (2007). IEEE 802.3 High Speed Study Group 100GbE Silicon Photonics Platform Considerations. IEEE 802.3 HSSG, Monterey, CA.

Kirchain, R. and F. Field (2000). Process-Based Cost Modeling: Understanding the Economics of Technical Decisions. Encyclopedia of Materials Science and Engineering. Oxford, U.K.

Kirchain, R. and L. Kimerling (2007). "A roadmap for nanophotonics." Nature Photonics **1**(6): 303-305.

Lebby, M. (2006). Optoelectronics: global markets and technology trends, Optoelectronics Industry Development Association (OIDA).

Liu, S. and E. Fuchs (2007). Envisioning the Future of Optoelectronic Interconnects: Production Economics of InP & Si Platforms for 100G Ethernet LAN Transceivers. Communication Technology Roadmap Spring Meeting, 2007, "Road to High Volume Photonics." Cambridge, MA, MIT Microphotonic Consortium.

McCall, T. (2006). Gartner Says Semiconductor Industry Grew 11% with Worldwide Revenue Reaching $261 Billion in 2006 Stamford, Gartner.

Nanez, R. and A. Iturralde (1995). Development of cost of ownership modeling at a semiconductor production facility. Advanced Semiconductor Manufacturing Conference and Workshop, 1995. ASMC 95 Proceedings. IEEE/SEMI 1995, IEEE.

O'Brien, D. and M. Schabel (2005). Microphotonics: Hardware for the Information Age: Next Generation Transceivers. Cambridge, The Microphotonics Center at MIT.

Ragona, S. (2001). "Cost of ownership." Photonics Spectra **35**(9): 12-12.

Saxena, V. (2007). Bandwidth drivers for 100 G Ethernet. IEEE 802.3 HSSG, Monterey, CA.

Schabel, M., E. Fuchs, et al. (2005). Microphotonics: Hardware for the Information Age: Current State of the Industry. Cambridge, The Microphotonics Center at MIT.

Schmitt, A. (2006, May 23). "Bookham, China, and the Optical Component Market." Retrieved October 1, 2001, from http://www.nyquistcapital.com/2006/05/23/bookham-china-and-the-optical-component-market.

Schmitt, A. (2007, June 8). "Infinera-The Optical Component Company That Wasn't." Retrieved March 13, 2008, from http://www.nyquistcapital.com/2007/06/08/infinera-the-optical-component-company-that-wasnt/.

SEMATECH. (2008). "SEMATECH History." Retrieved December, 2007, from http://www.sematech.org/corporate/history.htm.

Si Technology Working Group (2008). Electronic-Photonic Markets, Prospects for Standardization, and the Consequences of Energy: A Roadmap Progress Report MIT Microphotonics Center Spring Meeting, Cambridge, MA.

Sikorski, S., R. Krueger, et al. (1989). "A systems approach to the evaluation of packaging design alternatives." Int. J. Hybrid Microelectron. **12**(1).

Wirbel, L. (2006). "Long Way from Lab to Real Life for 100G." Electronic Engineering Times.

Wright, R. (2001). Cost Resource Model. S. Liu. Cambridge.

Zhang, Y. E. (2004). Next Generation Optical Receivers: Integration and New Materials Platform. Department of Materials Science and Engineering. Cambridge, MIT. Master of Engineering in Materials Science.

Appendix A. Background Information on IEEE HSSG

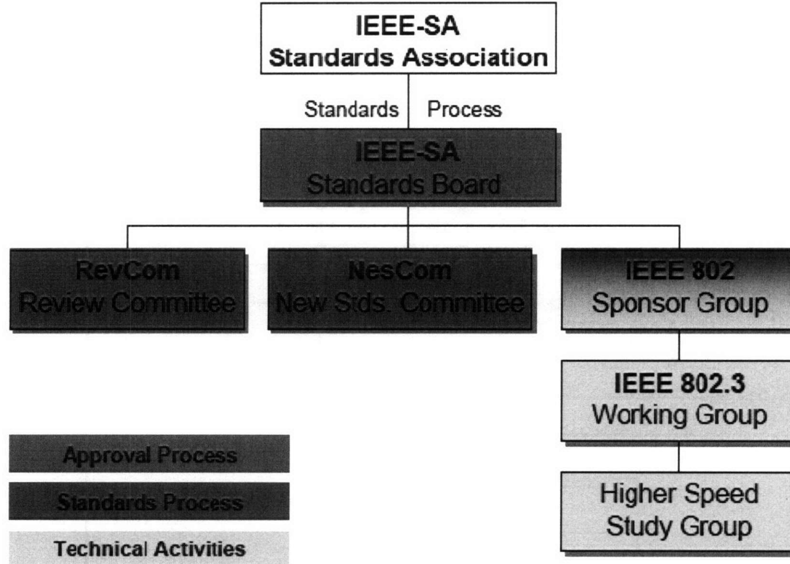


Figure 44. IEEE Structure
(HSSG 2007)

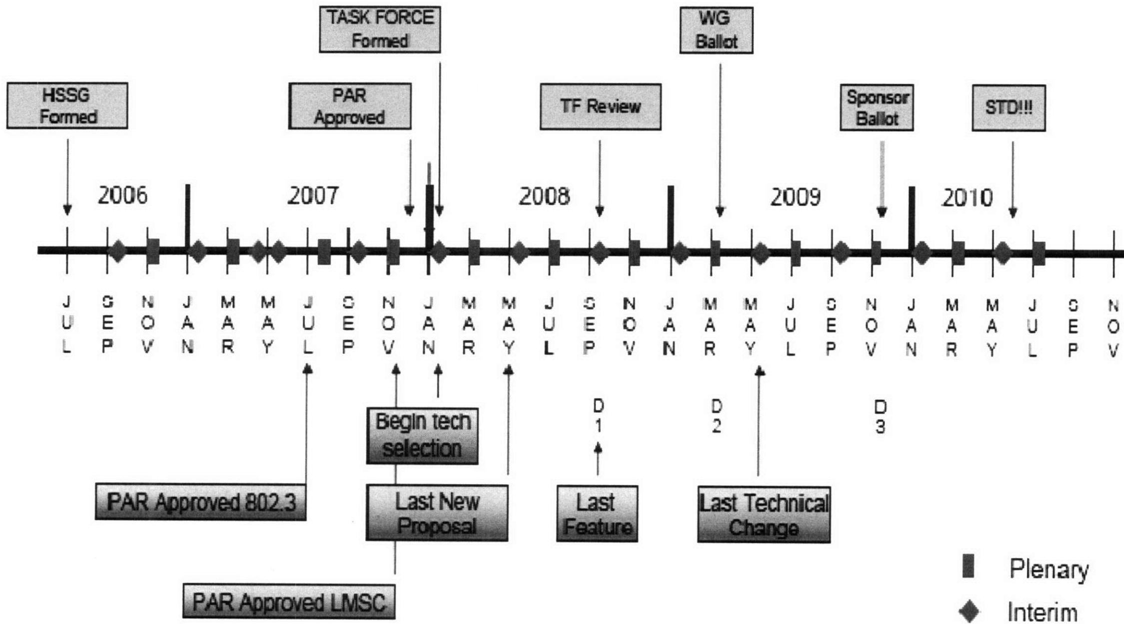


Figure 45. IEEE HSSG 100GE Timeline
(HSSG 2007)

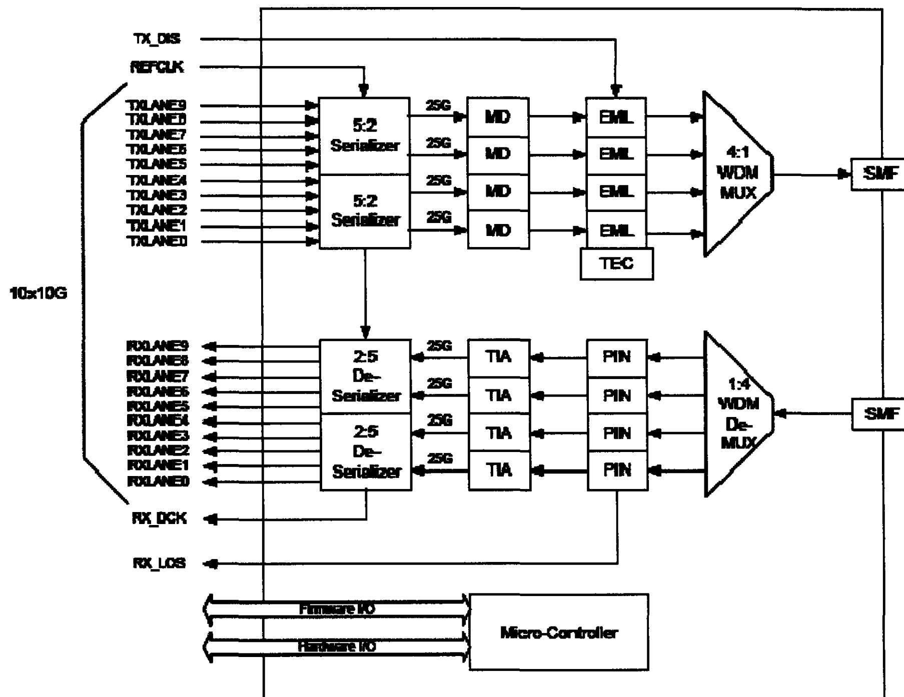


Figure 46. 4x25G Transceiver Architecture for Metro Applications includes SOA(s) (Cole and Huebner 2007)

Appendix B. Optical Components – Computing Market Segment Questionnaires

The Microphotonics Consortium at M.I.T. is conducting research on the adoption of optoelectronic components in high performance computing market segments. As an element of this research, this interview attempts to ascertain the current and future requirements for interconnects used in each of the four market segments (HPC Box-to-Box, Backplane-to-Board, Board-to-Chip, Chip-to-Chip). Please provide responses for **TWO** of the four market segments on this page. Also, please provide responses to the **TWO** corresponding questionnaires.

Market / Technology Attributes	HPC Box-to-Box	Backplane-to-Board	Board-to-Chip	Chip-to-Chip
Bit Error Rate Response Type: Rate Range: 10 E-3 to 10 E-20	10E-12	10E-12	10E-12	10E-15
Wavelength Response Type: nanometers Range: 400 nm to 1550 nm	850 nm	850 nm	850 nm	850 nm
Link Types Response Types: Freespace, Waveguide, POF to MM GOF to SM GOF, OTHER	MM GOF	MM GOF	MM GOF	MM GOF
Link Lengths Response Type: mm to Km	1m – 100 m	<1m to 2m	<1m	<1m
Temperature Response Type: Degrees Celsius Range: -55 C to 150 C	Controlled environment ~ RT	Controlled environment ~ RT	Controlled environment ~ RT	Controlled environment ~ RT
Standards Response Type: ITU, IEEE, OIF, Others	IEEE HSSG, IEC	IEEE HSSG, IEC	IEEE HSSG, IEC	IEEE HSSG, IEC
Reliability Response Type: Failures in Time	10	10	3-5	.3

Suggested Definitions
Bandwidth Demand = Mbps per port

HPC Box-to-Box Scenario

Define **link** (We assume that a link is a dedicated transmit or receive signal on a single fiber).

Data Transceiver Properties	2007	2010	2013	2016
Data Rate Per Link Response Type: Gbps Response Range: 5-100 Gbps				
Energy Density Per Link Response Type: Watts per Gbps Response Range: 5 mW – 50 mW **Supply Power Per Link** Response Type: Watts Per Link Response Range: 0W – 5W				
Optical Connector Type Response Type: Direct to LC to SNAP 12				
Electrical Connector Type Response Type: Differential LVDS, SFP, XAUI, XFI, SFP+				
What cost would need to be achieved by the time of adoption? Response Type: $ per Gbps Response Range: $0.50 - $5.00				
Link Properties	**2007**	**2010**	**2013**	**2016**
Number of Links (link = 1 fiber) Response Type: Per blade or drawer Response Range: 100-10000				
Channels per Link (wavelengths/fiber) Response Type: Ribbon fibers or Wavelengths				
Physical and Economic Properties	**2007**	**2010**	**2013**	**2016**
# fiber/sq. cm Response Type: # fibers Response Range: 1 - 96				
Transceiver on PWB or Active cable assembly? Response Type: PWB or ACA				
Cost Target for an entire Link (end to end) Response Type: $ per individual link Response Range: 50 - 300				
# links/Transceiver Response Type: # Response Range: 1- 96				
Estimated Annual Unit Volumes of Transceivers Response Type: In terms of Transceiver Volumes Response Range: 0 – N				

Appendix C. Benchmarking with External Models

As discussed in Chapter 3, COO is the established modeling method for the semiconductor industry. Since the Si fab modeling capability is new in the CTR FOP-PBCM, results are benchmarked with two well-recognized industry models developed based on COO: SEMATECH's Cost Resource Model (CRM) and IC Knowledge's IC Cost Model. Assuming the same global parameters in the external models, we break down the Si Two Chip frontend process flows into a list of equipment groups and input them into the two models. Results are shown in Table 23.

Table 23. Three model results comparison for Design 4 in an 8 inch Si fab at 20K wafers start per month

	PBCM	SEMATECH	IC Knowledge
Material	$85	$80	$38
Wafer	$291	$294	$294
Direct Labor	$47	$64	$53
Energy	$1	$1	N/A
Machine/Equipment	$390	$393	$402
Building/Maintenance/ Overhead/Indirect Labor, space	$187	$255	$174
Cumulative Yield	0.85	0.85	0.85
Depreciation years[17]	5	5	5
Processed Wafer Cost	**$1,001**	**$1,086**	**$961**

Processed wafer costs are fairly similar across the FOP-PBCM, CRM, and IC Cost models in all major cost categories. This benchmarking analysis suggests that the CTR's FOP-PBCM provides reasonably accurate results.

[17] Assume a zero discount rate in PBCM for this comparison

Appendix D. Analysis on Wafer Size Transition

The semiconductor industry keeps true to its creed, "smaller, faster, and cheaper," by having maintained an annual growth rate of 15% in revenue for the past 30 years (England and England 1998). This industry has also been able to increase its productivity by 25-30% per year through a combination of wafer size transitions, shrinking device geometries, equipment upgrades, and yield improvements (ICEC 1997). The last three factors are all accounted for in the cost modeling results shown in Chapter 4. This section studies the effect of wafer size transition on production cost for the Si material platform. Figure 47 shows a roadmap for each successive wafer size transition.

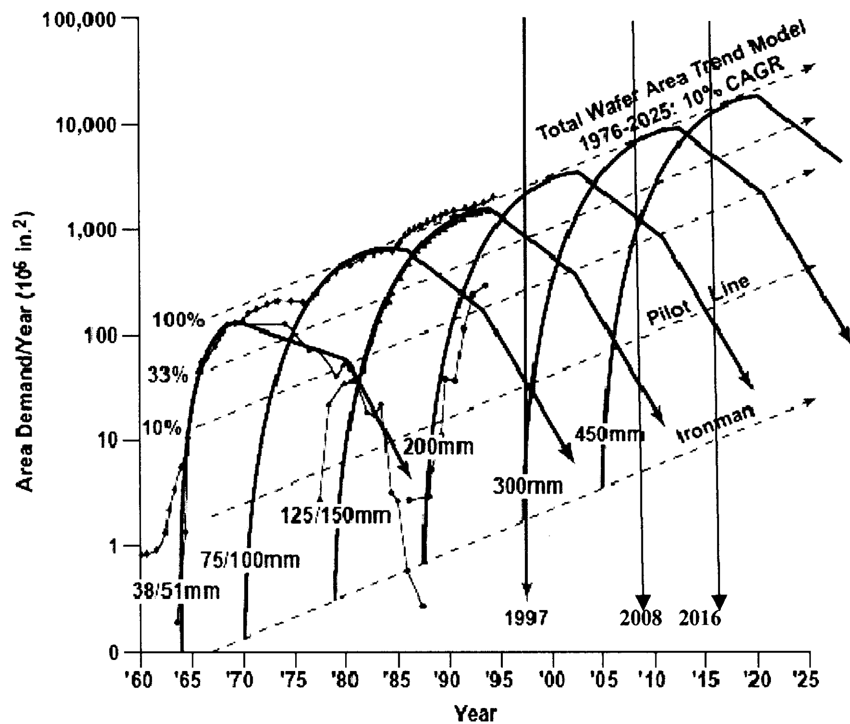


Figure 47. Wafer size lifecycle
(ICEC 1997)

This figure indicates that in year 2008, the industry is at the peak of 300mm fab utilization and at the downward slope for 200mm fab utilization. For Si photonics to take advantage of existing CMOS fabrication infrastructure, one key question is whether Si photonics should be build on 200mm or 300mm lines. 200mm fabs are currently being phased out for IC production and could be a starting platform for electronic-photonic integrated IC. This research explores the cost competitiveness of fabricating the Hybrid and Si Two Chip designs on a 200mm verses

300mm fab. Different assumptions on die size, number of die per wafer, and substrate cost are used in the model (Table 24). The FOP-PBCM allows users to enter a global throughput factor used for the majority of processes in the model (lithography is excluded), defined as a ratio of 300mm over 200mm machine throughputs. Users can experiment with a variety of throughput factors to get quick cost estimates for producing the same design on a 300mm line. Results in this study are based on a throughput factor of 1, which is an optimistic estimate assuming the majority of 300mm machine throughputs are on par with 200mm machine throughputs (excluding lithography) (Beals 2007).

Table 24. Frontend model assumptions for 200mm and 300mm wafer size

	Si Two Chip (300mm)	**Si Two Chip (200mm)**	**Hybrid (300mm)**	**Hybrid (200mm)**
Die Size	14 mm x 15 mm		12 mm x 15 mm	
Die/Wafer	210	90	245	105
Material	SOI			
SOI Wafer Cost	$800	$250	$800	$250
Machine Life	5 years			
Cumulative Yield	85%		90%	

SEMATECH and other industry experts estimate that for the 200mm to 300mm wafer transition there is a 125% increase in wafer area, a 20% to 40% increase in equipment cost, a 20% to 40% decrease in tool throughput (wafer/hour) and a 30% cost reduction per unit of starting wafer. Labor, material, and emissions should be comparable between the two wafer sizes. The cost per die can be 25% to 40% less in a 300mm fab for microelectronic components (ICEC 1997).

Conclusions drawn from the PBCM model result are less optimistic for the photonic chips. The result suggests that photonic chips should be manufactured in 200mm fab for both Hybrid and Si Two Chip designs. 300mm fab does not provide per unit cost savings until reaching a volume near 3 million APV (Figure 48 and Figure 49). The cost saving stabilizes at around $4 per die between the two designs after 10 million APV.

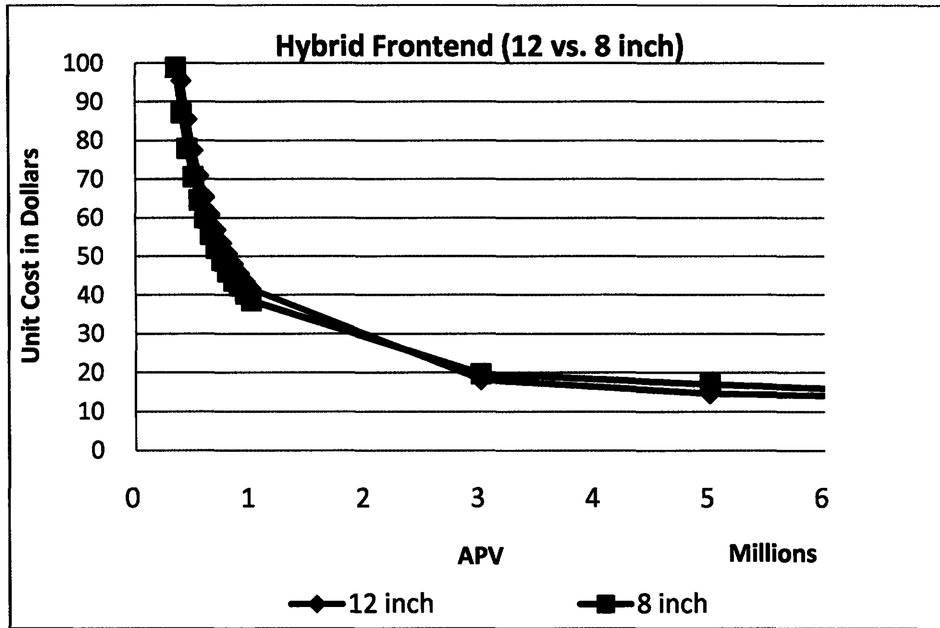


Figure 48. Hybrid frontend cost comparison, 200 mm vs. 300 mm

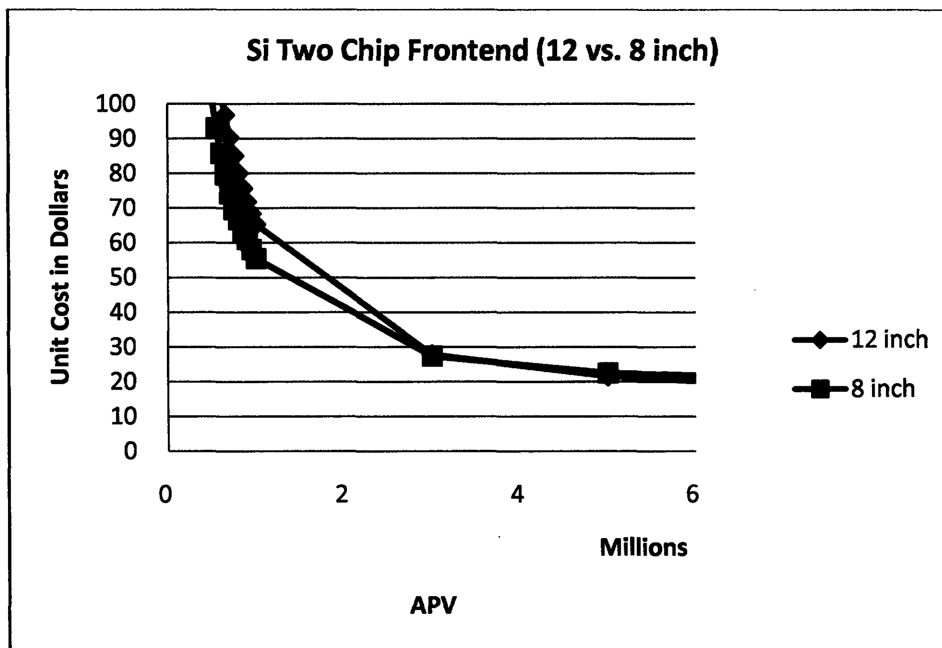


Figure 49. Si Two Chip frontend cost comparison, 200 mm vs. 300 mm

Figure 50 displays two pie charts comparing the per-die cost of the Si Two Chip design for the two wafer sizes. Die cost for the 300mm process is 81% of the die cost for the 200mm process, indicating a 19% cost saving for this wafer transition. A 19% cost saving is lower than the 25% to 30% industry average (ICEC 1997). The lower percentage of cost savings is due to the large increase in SOI substrate cost from the 200mm to 300mm size upgrade. If, instead of

$800 per 300mm wafer, $400 is assumed as the new substrate cost[18], the percentage of cost savings is increased to 31%, bearing in mind that a high throughput factor is used for the 300mm process.

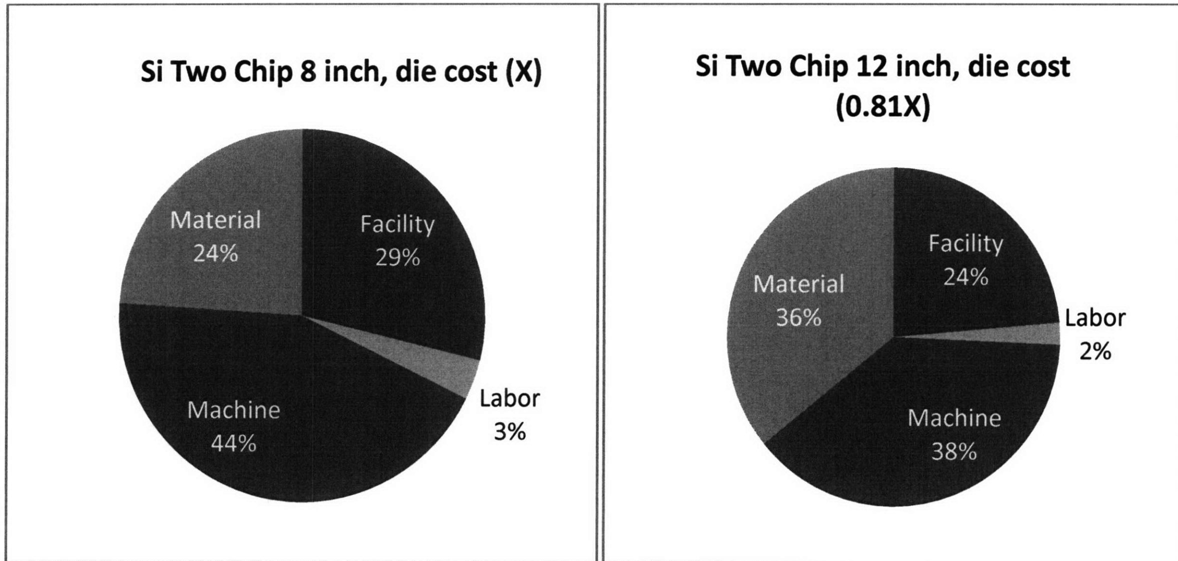


Figure 50. a) Si Two Chip 200mm wafer die cost, total cost: $17.5, b) Si Two Chip 300mm wafer die cost, total cost: $14.2 at 18.55M APV

[18] $250 x 2.25 (wafer area increase) x 0.7 (per unit of material savings in 300mm) = $400