Industry Dynamics within Semiconductor Value Chain

IDM, Foundry and Fabless

by

Shih-Lung Wu
M.S. Electrical Engineering
University of Southern California

Submitted to the Alfred P. Sloan School of Management in Partial Fulfillment of the Requirement of the Degree of

Master of Science in Management of Technology

at the

Massachusetts Institute of Technology

May 2003

© 2003 Shih-Lung Wu. All rights reserved.

The author hereby grants to M.I.T. permission to reproduce and to distribute, publicly, paper and electronic copies of this thesis document in whole or in part.

Author

Sloan School of Management
May 15, 2003

Certified by

Charles Fine
Chrysler LFM Professor of Management
Thesis Supervisor

Accepted by

David Weber
Director, Management of Technology Program
Industry Dynamics within Semiconductor Value Chain

IDM, Foundry and Fabless

by

Shih-Lung Wu

Submitted to the Alfred P. Sloan School of Management in
Partial Fulfillment of the Requirement of the Degree of
Master of Science in Management of Technology

May 8, 2003

ABSTRACT

This thesis intends to analyze the semiconductor value chain and identify critical factors, which impact the future industry structure. The main focus is in the dynamics among IDM (Integrated Design Manufacture), foundry and fabless companies. This analysis will utilize double helix [1] model with consideration of economics of scale, fixed cost and product life cycle. Furthermore, the cost and time-to-market factors involved in SoC (system-on-chip) and FPGA/PLD development will be explored. Based on the analysis, a tool will be developed to evaluate make-or-buy (developing own SoC or using off-shelf FPGA/PLD) decision.

Thesis Supervisor: Charles Fine
Title: Chrysler LFM Professor of Management, Sloan School of Management
Acknowledgements

I dedicate this thesis to the memory of my father, Ching-Shyang Wu. He taught me how to prepare and get most out of the bad time. I learned the economics S-curve from my father since I was a teenager. There are always people making money during recession. There are even more people losing money during the boom time. How one adapts change and prepares for the tipping point is the key to success. By the same token, how one is patient to deal with bottleneck is the key to have any breakthrough. This is a great asset I inherited from my father.

I thank my mother, Su-Ching Wu, my wife, Grace and my daughter Serena for their unlimited support during my study in MIT. As a son, husband and father, I appreciate their understanding for my absence during past twelve months.

I also thank Professor Charles Fine for his insight and guidance in framing the structure of this thesis. Without his double-helix framework, this thesis will lack the foundation to dig into the issue.

I thank my friends, Benjamin Tandiono for his help in data collection and Ding Li for his information leading me to join this program.

Last but not least, I have to thank again to my wife, Grace for her encouragement to satisfy my curiosity. Few spouses will agree their significant half to spend a hundred thousand dollars and one-year time just because he/she is curious about how economics and business work. This is the only reason I could come out with a year ago. Amazingly, she agreed.
Table of Contents

1 INTRODUCTION ........................................................................................................ 5
  1.1 OVERVIEW OF SEMICONDUCTOR VALUE CHAIN ............................................. 6
  1.1.1 Equipment........................................................................................................... 8
  1.1.2 Packaging and Testing....................................................................................... 10
  1.1.3 EDA..................................................................................................................... 12
  1.1.4 SIP (Silicon Intellect Property)........................................................................... 14
  1.1.5 IDM (Integrated Design Manufacture)............................................................ 15
  1.1.6 Fabless & Foundry......................................................................................... 17

2 SEMICONDUCTOR INDUSTRY DYNAMICS .......... 19
  2.1 FORCES DRIVING THE STRUCTURE CHAGE .................................................. 20
  2.1.1 Forces of disintegration push toward a horizontal configuration .................... 20
  2.1.2 Forces of integration push toward a vertical configuration ............................. 27
  2.2 CONCLUSION...................................................................................................... 30

3 THE STATE OF SOC MARKET................................. 33
  3.1 THE BENEFIT OF SOC........................................................................................ 33
  3.2 NEW ENTRANT AND ENTRY BARRIER................................................................. 35
  3.3 CUSTOMER INFLUENCE....................................................................................... 39
  3.4 COMPETITION ..................................................................................................... 40
  3.5 FOUNDRY'S INFLUENCE ..................................................................................... 41
  3.6 SUBSTITUTE.......................................................................................................... 41

4 THE IMPACT OF FPGA/PLD TO SOC INDUSTRY .. 44
  4.1 NEW DEVELOPMENT IN FPGA/PLD .............................................................. 44
  4.2 SOC or SoRC: A QUESTION OF TOTAL COST ............................................... 45
  4.3 CONCLUSION...................................................................................................... 57

5 SUMMARY .............................................................................................................. 59
1 Introduction

In this thesis, semiconductor industry is defined as "companies, which design and/or manufacture the chip", therefore it includes only IDM (Integrated Design Manufacture), foundry and fabless companies. Semiconductor value chain is all of technologies/industries required to produce a chip or IC (Integrated Circuit). Figure 1.1 shows current industry business models.

Semiconductor Business Models

<table>
<thead>
<tr>
<th>IDM Model</th>
<th>Fabrite Model</th>
<th>Fabless Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing/</td>
<td>Marketing/</td>
<td>Marketing/</td>
</tr>
<tr>
<td>Sales</td>
<td>Sales</td>
<td>Sales</td>
</tr>
<tr>
<td>Design/IP</td>
<td>Design/IP</td>
<td>Design/IP</td>
</tr>
<tr>
<td>Systems</td>
<td>Systems</td>
<td>Systems</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>Manufacturing</td>
<td>Manufacturing</td>
</tr>
</tbody>
</table>

Companies: IBM, Intel, Texas Instruments
Pros: Control over their own roadmaps
Cons: Cost, risk, swings in utilization
Prerequisite: Must be a $7B+ to support aggressive manufacturing

Companies: Motorola, Infineon, ADI
Pros: Have some control over process technology, yet chance to have access to leading edge technology
Cons: Once decision is made to reduce or stop investment, ability to reverse is difficult.

Companies: Xilinx, Nvidia (over 600)
Pros: Ability to focus on value added, while having wide access to leading-edge technology. Upside during capacity shortages, no worry with utilization during oversupply
Cons: Dependence upon limited number of foundry partners

Source: Fabless Semiconductor Association

Figure 1.1

Since the transistor was invented in 1950s, semiconductor industry has come a long way in transforming its downstream industries, which include computer, control, consumer electronics and communications etc… Before 1990’s, semiconductor industry is dominated by IDMs (Integrated Design Manufacture), which not only does design and manufacture its own chips, but also design its own EDA (electronics design automation) tools, sometimes even designed and modified its own machine tools. Due to the scale of economy, boom-bust cycle and the productivity gain realized from Moore’s Law—A trend observed by Intel cofounder Gordon
Moore in 1965 in which the number of transistors in integrated circuits doubles every 18 months—semiconductor industry itself has been experiencing a major change in industry structure for last ten years. This change is mainly because of the emergence of pure-play semiconductor foundry and fabless firms. Current industry value chain is showed in figure 1.2. MEMS (micro electrical mechanical system) and specialty foundry will not be discussed in this thesis.

![Figure 1.2](image)

1.1 Overview of Semiconductor Value Chain

Firms in semiconductor value chain can be divided into two groups, design and manufacturing. Those industries dealing with designing chips are SIP (Silicon Intellect Property), EDA (Electronic Design Automation), fabless and IDM (Integrated Design Manufacture). Material, equipment, packaging, testing, IDM and foundry are industries involved in manufacturing the chip.

Depends on the complexity of a chip and related design approach, designing a chip can either involve less than ten or more than a couple hundreds of VLSI engineers. The functionality of a
chip can be as complex as an Intel CPU (full custom ASIC) or as simple as small glue logic PLD (Programmable Logic Device). Chip design processes require sophisticated tools (EDA) and/or reusable 3rd party silicon intellect property (SIP). These tools and reusable blocks help engineers to tackle complex tasks difficult for human to handle.

Manufacturing a chip is a complex process involving hundreds of steps that must be executed with great precision. In simplified terms, Chip manufacturing involves three basic operations: deposition, patterning and etching. Many processing steps are repeated many times in IC or chip fabrication, resulting in the buildup of microscopically thin layers of materials. In the process of building these layers, thousands or millions of transistors are created (transistor cycle) and interconnected (copper interconnect cycle). When the process is complete, a single wafer will contain hundreds of individual chips that are then cut from the wafer, tested for their electrical properties, packaged and assembled. Figure 1.3 shows the manufacturing flow.
1.1.1 Equipment

Chips are made in specialized factories called "fabs," or fabrication plants. Often referred to as the most sophisticated manufacturing plants in the world, advanced fabs can cost anywhere between $1 billion and $2.5 billion to build and equip. Today, there are approximately 900 fabs [2] around the world producing millions of chips each day and billions of chips every year.
The chip manufacturing process in a fab occurs in a “cleanroom,” a special area where airborne contaminants are scrupulously controlled. In addition to the cleanest air, stringent cleanliness requirements are also followed in the use of chipmaking materials and processing equipment. Figure 1.4 shows different technologies used in the cleanroom.

![Cleanroom Technologies Diagram]

Of these cleanroom technologies, following three systems are considered to be the most critical and expensive.

**Deposition systems**

These systems deposit insulating and conducting film layers and can be divided into following categories.

- **Metal deposition system**

  Using Physical Vapor Deposition (PVD) argon atoms are shot at a “target” of pure metal. These metal atoms then chip off and deposit on the wafer surface.

- **Dielectric deposition system**
Using Chemical Vapor Deposition (CVD) technology, an insulating material is deposited on the wafer surface, forming a thin layer of solid material on the chip.

- Silicon and thermal deposition system

  Epitaxial Silicon ("epi") is a special layer of extremely pure silicon crystal grown on some wafers to enhance chip performance.

**Photolithography systems (or steppers)**

These systems print the chip’s design on the wafer.

**Etch systems**

These systems remove materials to create the chip’s structure and can be divided into three categories, dielectric etch, metal etch and silicon etch.

1.1.2 Packaging and Testing

**Packaging**

IC or chip packaging is the essential technology, connecting the bare chip to the PCB (Printed Circuit Board). It serves as a reliable and easily testable vehicle for susceptible chips. There are following different kind of packaging technologies.

- **BGA**

  BGA is the package which places its output pins in the form of a solder ball matrix. The traces of BGA are generally fabricated on laminated substrate (BT-based) by thick film process. This means that the entire area of substrate can be used to route the interconnection.

- **Multi-Chip Module (MCM)**

  MCMs can integrate different functional chips into a mini-substrate instead of inserting individual packages onto a large PCB. There are two different structures, 2D (side-by-side) or 3D (stacked) for dice placement. These structures are commonly used to combine a
complex processing chip with several cache memory chips. Both 2D and 3D topology can be incorporated to further reduce the packaging size.

- **Dual-in-Line Packages**
  Long-lasting IC packages in consumer products, auto devices, memory, and high frequency ICs, dual-in-line packages have evolved into state-of-the art ones owing to their robust reliability and great improvement on packaging capability. Wide usage of PTH (plated through hole) and SMT (surface mount technology) assembly, dual-in-line packages provide an assortment of packaging capability, especially in low pin count devices at a lower manufacturing cost.

- **QUAD Packages**
  Quad packages have been used for years for ASICs, DSPs, micro-controllers, and memory ICs. They offer an excellent low cost solution for moderate and low pins ICs.

- **CSP**
  Bare dies on CSPs (chip scale package or chip size package) occupy over 80% of the whole package. This makes CSPs not only easy to be handled, tested, and assembled, but also a substitute for known good die (KGD) as test and burn-in methods are still not available. CSPs are suitable for low or moderate pin count Ics such as memory, RFICs, and communication ICs.

**Testing**
Testing is the final gatekeeper of semiconductor manufacturing flow. Shipping a defected chip to customer carries huge liability. Not only does it damage firm’s reputation, but also carries a big price tag. Quite often, the cost to detect and repair a system product once the chip already mounted on the PCB are far more expensive than the cost of the chip, let alone shipping the
system product to end user then got returned. The major steps to test chip are wafer sort (chip probing), final test and burn in test. Figure 1.6 shows how test was done.

Wafer Sorting (CP):

![Diagram of wafer sorting process]

Final Test (FT):

![Diagram of final test process]

**Figure 1.6**

1.1.3 EDA

EDA stands for Electronic Design Automation. To understand the rapidly growing, almost four billion dollar EDA industry, it helps to define what it means for the words behind those three letters, "EDA":

**Electronic**

Electronic—from computer chips, cellular phones, pacemakers, controls for automobiles and satellites to the servers, routers and switches that run the Internet. Everything made by the nearly $1 trillion electronics industry results from designers using EDA tools and services. As electronics become even more complex and pervasive, the EDA industry is more vital to the
continued success of the global economy.

**Design**

This part of the development cycle is where creativity, new ideas, ingenuity and inspiration come to the fore. This is also where designers try to model the behavior of their designs and analyze the complex interactions of millions of constituent parts in their designs to ensure completeness, correctness and manufacturability of the final product. These steps are crucial, because it is impossibly difficult, expensive and time consuming to "build a chip first and fix it later". Because the designers in EDA industry are mostly electrical engineers ("hardware engineers") and computer scientists ("software engineers"), some segments of the EDA industry are also called, "Computer Aided Engineering" (CAE). EDA is also referred to as "Electronic Computer-Aided Design" (ECAD), acknowledging the crucial role EDA plays in the design phase.

**Automation**

Imagine the difference between designing a small house versus designing a mile-high skyscraper. For the skyscraper you need to design sophisticated structural, electrical, plumbing, security and environmental systems, communications and computer networks, elevators, etc. all working together. This is analogous to the dramatic increase in complexity that designers must tackle in electronics today. It is this complexity enabled by the relentless onslaught of Moore's Law drives the need for automation. For more than 30 years this has been the driving force behind the electronics revolution. Engineers need to validate their concepts, model and analyze their designs, identify and eliminate problems before making production commitments. EDA helps them get it done right.

To illustrate the importance of EDA tools during chip design, we use a SoC (system-on-chip) design flow as an example. As showed in figure 1.5, almost all major steps in the design flow require EDA tools. Together, these steps account for at least half of time of the development
cycle. Typically, it takes one and a half to two years to develop a chip. A good EDA tool can surely save lots of development cost and time.

1.1.4 SIP (Silicon Intellect Property)

Silicon Intellectual Property firms are also called chipless firms since they do not create chips but create intellectual property, i.e. designs that can be used by fabless firms or integrated players to create chips. Most large semiconductor companies have SIP and retain it in-house, but SIP firms create the IP and try to sell it or license it to other companies. This model involves a higher gestation period, and there is no guarantee of any return on investment. To follow the SIP model successfully, a firm needs deep technological expertise along with the required capital.
Similar to the products industry, SIP firms can have significant gains in the long term by way of revenues from licensing and royalty. Unlike the design services industry, this domain requires firms with enough working capital in order to survive the initial years. Some design service firms leverage the experience and effort involved in a project and subsequently spun off into an IP. While IP may not contribute significantly in the short term, it offers other advantages. For example, the same IP block can be used for a different design after some tweaking. This helps in ramping up a project in quick time.

There are significant differences between an IP block, an IP component and an IP core. An IP block is simply a design that has been developed by a firm and can be re-used and customized, but the IP block may lack of well structured interface logic. This disadvantage makes it difficult for chip companies to integrate the IP block with their own product. An IP component can be compared to a product where a company buys the component and integrates it into its own design. An IP core is a critical part of the entire design and requires considerable expertise. A client typically pays those SIP firms specialize in IP core space not only the licensing fee but also a share of revenues earned on the design. For example, processor core specialists ARM and MIPS even charge royalty on how many instances of the processor core designed in a chip.

1.1.5 IDM (Integrated Design Manufacture)

IDMs are those 800-hundred-pound gorillas in semiconductor industry space. These companies have their own chip manufacturing facility. In other words, they own fabs. As the name implied, IDM is an integrated firm conducting chip design, production, marketing, supply chain management and after sale support. The big name in this industry is company like Intel, Texas Instruments, Motorola, Infineon and Toshiba etc… Although foundry is getting momentum lately, IDMs still own more than 80% of worldwide semiconductor market. Figure 1.7 shows the market share of IDM and foundry.
Due to the huge economics of scale required to justify capital investment in building a new fab, IDM is gradually reducing their investment in fab. Following is some recent development.

- Infineon plans to outsource up to 50%, up from 10% today.
- Motorola plans to outsource up to 50% of its production.
- Texas Instruments stated an outsourcing goal of 40% over the next several years.
- Toshiba plans to boost outsourced chip production to 40% from 7% in 1999-2000.
- National Semiconductor currently outsources 20% of its production.
- LSI Logic currently outsources 30% to 35%, and is forecast to outsource more than 60% in the next few years.

According to a research report conducted by Joe Osha of Merrill Lynch (March 2002), most of second and third-tier IDM will be forced to adopt a pure fabless strategy or a fablite strategy to remain viable. Over the next 18 to 24 months, up to 20 IDM will make this transition.
1.1.6 Fabless & Foundry

Fabless and foundry have the strong tie and interest to see each other prosper. Before late 1980s, fabless firms always lived under the shadow of IDMs. On the one hand, fabless try to grow in the new or niche market, but as soon as market demand began to take off, IDM will soon step in and compete. On the other hand, fabless depended on IDM for the manufacturing capacity, since fabless can’t afford to build its own fab. To fabless firms, IDM is the worst supplier and competitor. With the emergence of foundry business model, the separation of chip design and manufacturing functions becomes feasible. Fabless firms no longer have to rely on IDMs for manufacturing capacity. The rise of the fabless model was further reinforced by two reasons. First, the emergence of new semiconductor applications (networking and communications etc...) and rapidly changing technology introduced risks make IDMs reluctant to step in new market during the early stages. Since new applications tend not to have dominant design [3], the total addressable market size is small in the beginning. Products with small market potential can not justify IDM’s high capital investment in the fab, therefore IDM usually wait until the market grow big enough. By the time, market becomes viable, IDM normally miss the market window due to the long product development cycle. In terms of the risk involved in changing technology, smaller design groups are naturally able to re-deploy their resources more efficiently to cope with such market risks. This, coupled with a reduction in transaction costs associated with information transfer from design to manufacturing, increased the attractiveness of firms to specialize in semiconductor design. Second, increasing economies of scale in manufacturing arising from higher manufacturing investment requirements meant that risks associated with fluctuating market demand for semiconductors could be mitigated by aggregating demand from a number of smaller fabless firms. This naturally led to the emergence foundries that concentrate solely on semiconductor manufacturing. The attractiveness of the “fabless-foundry” model was
self-perpetuating in that it mitigated the risks faced by design firms of losing their product marketing know-how to contract manufacturers, thereby strengthening the emergence of the fabless-foundry model. In essence, this new business model enhances the efficiency of the semiconductor industry by effectively apportioning the risk in technology and market forces to firms that are best suited to diversify and mitigate the costs of these risks.

Table 1.1 shows the share of foundry as percentage of worldwide semiconductor wafer production.

<table>
<thead>
<tr>
<th>Year</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Wafers (K)</td>
<td>93,132</td>
<td>103,989</td>
<td>107,374</td>
<td>117,607</td>
<td>135,177</td>
<td>155,832</td>
</tr>
<tr>
<td>Foundry Wafers (K)</td>
<td>12,912</td>
<td>15,143</td>
<td>16,826</td>
<td>21,399</td>
<td>27,045</td>
<td>31,137</td>
</tr>
<tr>
<td>Foundry Market Share</td>
<td>13.9%</td>
<td>14.6%</td>
<td>15.7%</td>
<td>18.2%</td>
<td>20.0%</td>
<td>20.0%</td>
</tr>
</tbody>
</table>

Table 1.1 Source: Fabless Semiconductor Association

Figure 1.8 shows the top 10 fabless revenue growth. Note, in 2001 quite a few of fabless firms actually have negative growth rate, but comparing to overall semiconductor industry, fabless firms actually fare better.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>QUALCOMM (QCT Div.)</td>
<td>$1.30B</td>
<td>$1.22B</td>
<td>$1.39B</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>NVIDIA</td>
<td>$375M</td>
<td>$735M</td>
<td>$1.37B</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Xilinx</td>
<td>$898M</td>
<td>$1.56B</td>
<td>$1.15B</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Broadcom</td>
<td>$518M</td>
<td>$1.13B</td>
<td>$961M</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Altera</td>
<td>$837M</td>
<td>$1.38B</td>
<td>$838M</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>Cirrus Logic</td>
<td>$532M</td>
<td>$739M</td>
<td>$534M</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>Adaptec</td>
<td>$776M</td>
<td>$768M</td>
<td>$422M</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>SanDisk</td>
<td>$247M</td>
<td>$602M</td>
<td>$366M</td>
</tr>
<tr>
<td>9</td>
<td>N/A</td>
<td>QLogic</td>
<td>$158M</td>
<td>$223M</td>
<td>$358M</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>PMC-Sierra</td>
<td>$262M</td>
<td>$695M</td>
<td>$323M</td>
</tr>
</tbody>
</table>
2 Semiconductor Industry Dynamics

Semiconductor industry is in transition now. This is the first time in history, the famous "Moore's Law" - defined as the density of transistor will be double every 18 months -, may have to be modified. Because of the difficulty in improving manufacturing process technologies, the period of doubling transistor density may have to be extended. In SIA's term, the presence of a potential "Red Brick Wall" or "100 nm Wall" (as indicated by the red cells in the technology requirements) that, by as early as 2003 [4], could block further scaling as predicted by Moore's Law. This chapter will take advantage of double helix model as the tool to analyze the dynamics of semiconductor industry structure. Those example illustrated in ClockSpeed [1] tends to be industries producing assembled products like computer, bicycle and automobile. The example showed here will demonstrate that the double helix mode is also applicable to industry with an integrated product, a silicon chip. In the end of this chapter, conclusion will be drawn to infer the possible future direction of semiconductor industry structure.

THE DOUBLE HELIX

Figure 2.1 Source: ClockSpeed, Charles H. Fine
2.1 Forces Driving the Structure Change

To predict the change of industry structure is like sailing a boat in the storm without a compass. This is because manufacturing a modern product normally required multiple technologies. Every technology may also required support of complement technologies. In addition, every product/technology has its own S-curve [5], therefore the end product/technology is the aggregation of all individual S-curve. Double helix model characterizes the industry change by classifying forces causing the industry change into two categories, forces of disintegration pushing toward a horizontal configuration and forces of integration pushing toward a vertical configuration. Following sections will analyze those forces behind semiconductor industry structure changes.

2.1.1 Forces of disintegration push toward a horizontal configuration

2.1.1.1 Entry of niche competitors

Due to the separation of semiconductor manufacturing and design, the initial fixed cost of starting a fabless start up is getting lower. This help to bring down the entry barrier. In addition, the value chain is moving toward to horizontal and specialization now. First, the EDA design tools are much more powerful and standardized now. Second, there are lots of SIP vendors providing various building blocks, which enable chip designer to integrate peripheral functionality with in-house design without the need to start from scratch. The work used to take a team of tens VLSI engineers in designing a chip can now be done in less than ten. Third, foundries are keen to provide extra services, like shuttle run -share photo mask cost by putting different test chip design from multiple customers on a same wafer- and low-cost IP library. This greatly reduces the initial investment required by fabless companies. For example, United Microelectronics Corp. and researcher IMEC announced the signing of an agreement between the No. 2 foundry and IMEC's Europractice IC Service offering access to UMC's Silicon
ShuttleR Multi Project Wafer platform for prototyping and production services [6]. The two are taking aim at European-based organizations with low-volume manufacturing needs, such as universities, research centers, start-ups and small niche companies, opposed to powerhouse customers. In the past, smaller companies sometimes had difficulty getting continued access to foundries. This collaboration shows that foundry is willing to provide a consistent solution for these smaller companies and is committed to help start-ups and fabless companies reach maturity.

As the entry barrier reduced, the niche fables firms flock in. Lots of these new fabless specialize in emerging product segment like home networking or multimedia application etc... In the meantime, the number of members in Fabless Semiconductor Association has been growing from 40 in 1994 to 315 in 2001. Figure 2.2 shows the revenue growth rate of fabless companies:

**Fabless Revenue Growth**

![Fabless Revenue Growth Chart]

*Figure 2.2  
Source: FSA and public information*
2.1.1.2 Complexity Involved in Integration

Semiconductor is such a complex product required sub-micron precision. All complementary technologies have to advance in the same pace. For example, advancement in material and EDA tools alone will not make any difference, if the lithograph or etching machinery can not take advantage of them. The technology and expertise involved in manufacturing a chip is completely different from that of required in designing chips. Integrating these two distinct activities into one firm may increase following complexity.

**Technology Complexity**

To achieve transistor density defined by Moore’s Law, SIA(Semiconductor Industry Association) has to setup several annual events to coordinate firms in semiconductor value chain and produce an annual ITRS(International Technology Roadmap for Semiconductors) report. This report discuss all of potential problems will be encountered by entire value chain in order to advance into next-generation semiconductor manufacturing process. Every complementary technology industry then try to work out a solution. The increasingly specialization in every element of technology value chain makes it difficult for IDM to maintain the vertical integration. For example, in terms of EDA tools, big IDMs used to design own tools set. Due to the complexity involved in SoC design verification and small-node (0.13um and 90nm etc...) timing/crosstalk checking, IDMs are having difficulty to add value by bringing EDA tool design in house. They all begin to use 3rd party EDA tools now. Following examples show the change of attitude:

- Silicon Metrics’ SiliconSmart product line will be implemented as part of Agere’s 90nm AGR90 ASIC platform. This is a common path to avoid adding 90nm capabilities to proprietary systems, according to Silicon Metrics, which also said IC design vendors such as
Agere are choosing to replace their proprietary system with commercially developed solutions [7].

- Cadence Design Systems Inc., an EDA tool vendor announced that IBM Corp. is going to buy more of its software and that the two are going to collaborate on their 90nm design flow. This deal is significant because IBM still develops a considerable amount of its own EDA software for the thorniest design issues, just like Intel Corp. IBM has said it is working diligently to get a return on that engineering investment [8].

In the area of chip design, many semiconductor companies— both IDM and fabless— are increasingly relying on external sources of the technical expertise for various components of the SOC design. The use of proven third-party SIP components allows semiconductor companies to meet market pressures while continuing to focus on the components of the SOC that constitute their core competencies. The complexity is forcing industry to look for standard components and 3rd party SIP vendors.

**Complexity in planning fab capacity**

Semiconductor industry has been experienced serious bullwhip effect [9] in its supply chain over the years. Considering the variability exists in semiconductor supply chain:

- **Demand forecast**: the boom-bust cycle shows the difficulty in forecasting demand.
- **Lead time**: the manufacturing lead time is long, at least 6 ~ 8 weeks.
- **Batch process**: the product usually manufactured and ordered in batch.
- **Inflating order**: due to the long lead time and big spike in downstream demand, inflating order is a common practice in this supply chain.

Because of above reasons, IDMs constantly worry about the big spike in under or over capacity. Due to the difference in business model, they can’t smooth the spike by aggregating multiple
external customers either. IDM only builds their fab to serve its own internal design team and product lines. Figure 2.3 shows the severity of boom-bust cycle.

![Graph showing annual growth rate over years](image)

**Figure 2.3 Source: Converge**

**Complexity of arranging financial resources to build a fab**

The cost of running a fab is skyrocketing. Raising billions of dollars every 18 months is a big challenge for the CFO in every IDM. The compound effect of high leverage and the expected ROI sometimes looks like a poison pill. Some bankruptcy cases of Korea conglomerates during Asia financial trouble in 1997 are largely due to the high leverage of their semiconductor arms. Figure 2.4 shows the estimated cost to own a fab.
The Cost of Building a Fab

![Bar Chart: Cost of Building a Fab](source: Dataquest and Reuters)

**Complexity in resources allocation**

Building a fab will take resources away from semiconductor firm's most value-add function, designing and marketing chip. To most IDMs, the economic rent provided by manufacturing chip is much less than designing chip, therefore even IDMs is outsourcing now. Figure 2.5 shows the growing trend of IDM outsourcing activities.

*Increasing IDM Outsourcing*

![Bar Chart: Increasing IDM Outsourcing](source: Dataquest and Reuters)
2.1.1.3 Organizational Rigidities

Due to the high initial investment in fab, the business model of most IDM is to design and manufacture chips with high volume potential, therefore they can best utilize the valuable fixed asset, the fab. A case in point is that most IDMs focus on general-purpose microprocessor, DSP, PC peripheral chipset and memory etc… In this business model, design or manufacturing low-to-mid volume products are not viewed as a viable business. This organizational thinking often makes IDMs ignore the emerging market. Even in some emerging market segment, IDM did spot opportunity and develop good product, but they tend to stick with the original design, since the market size is still not growing big enough for them to recoup existing investment. This attitude keeps IDM from continue investing in the follow up product and eventually lose in climbing the S-Curve. Another case is to deal with demand fluctuation. Even some IDM insiders recognized the need and opportunity to smooth demand of fab capacity, but the organization rigidity filtered or blocked the idea. The example is that Morris Chang quit Texas Instrument and found the first foundry, TSMC. Foundry acts as an aggregator to manufacture chips design by other semiconductor companies. The creation of foundry eventually enables hundreds of fabless firms to compete with IDMs.

Recognizing that foundry is a viable business model and its own organization structure is not set up to serve external customers, even IBM is taking action to transform its foundry business. For example, IBM Microelectronics announced its intention to focus on pure-play foundry services and establish a services business. IBM set up a new services business established to assist chip companies developing everything from complex chips to entire systems. IBM Engineering & Technology Services will offer design, foundry and consulting services, granting access to IBM’s library of IP.
2.1.2 Forces of integration push toward a vertical configuration

2.1.2.1 Technical advantage result from integration

**New process technology**

For low-k dielectric films at 130nm and copper process etc..., IDM clearly has an edge over foundry in terms of yield rate [10]. For example, the manufacturing issue in new process and yield rate forces UMC to closely aligned with Infineon and IBM for future process development. TSMC also has locked arms with Philips, ST Microelectronics, AMD and Motorola. What could emerge from this inflection point at 130nm is what industry historians might call “captive foundries”, meaning in the worst case, foundry may end up only produce less advanced product and become manufacturing facility for IDM to adjust capacity.

**Proprietary design tools**

Companies using third-party design tools failed to account for thermal migration issues because they were mainly looking for the electrical migration issues that aluminum taught them to monitor. Companies with their own internal chip design team and tools had the ability to model designs with thermal migration issues in mind and could work through them faster.

**Advanced process IP results from integration**

As the linkage between process technology and design becomes increasingly critical, the market is transforming from a traditional ASIC business to a complex design, co-development service with much higher levels of interaction and collaboration. "There continues to be a thriving ASIC market for the few companies that can build these large, highly complex products where the old methodology of designing it and then throwing it over the wall to manufacturing no longer works," said Jerry Worchel, senior analyst at the market researcher, in a statement [11]. With a wide variety of high quality IP, systems expertise and leading-edge process technology, IDM are well suited to serve this segment. For example, IBM and Toshiba announced their foundry
services for 90nm could very well be interpreted that some foundry customers are unhappy with the results they achieved from the pure-play foundries at 130nm—and they are counting on IBM and Toshiba to come through at 90nm. As a result, some High-end fabless like Xilinx begins to use two foundry partners that have different processes. The company uses IBM primarily as a technology development partner and UMC as its primary volume production partner. By the same token, IBM (as a technology partner) announces that it will team up with Charters (volume production partner) to provide foundry service [12].

2.1.2.2 Market power in bundling subsystems

There is a well known fact that companies with dominant position in key components are likely to bundle subsystems to extract extra consumer surplus. This is especially evident in IDM business. For example, Intel bundles its CPU with associated chipset, Motorola integrated its PowerPC core with other peripheral functionality and TI bundle its DSP with other chipset. On the one hand, bundling is one form of one-stop shopping. It can simplify hardware system design and accelerate customer's time to market. On the other hand, customers may end up be forced to use an inferior solution. A case in point is that Intel bundles 802.11b solution with its new wireless notebook CPU, Centrino. Although Intel's 802.11b solution lags behind 3rd party wireless products in performance and reach, customers are forced to use Intel solution.

An example to illustrate how IDM can exercise its power in bundling manufacturing and design can be seen in following announcement. Intel announced plans to fabricate a single silicon chip that combines both digital and analog functions utilizing SiGe process [13]. The new chip, which Intel expects to be on the market in 2004, should significantly lower Intel's manufacturing costs. SiGe process has great advantage in noise tolerance and higher bandwidth, which may be good enough to compensate inferior chip design, therefore rival fabless product using CMOS foundry may not be able to compete. Intel is well positioned because its recent announcement
builds on what industry observers see as two of the company's key strengths. Firstly, the company already leads the market for microprocessors, which provide PCs with computational power. In addition, Intel is the top provider of a key component used in the manufacture of cell phones, flash memory.

2.1.2.3 Profitability from proprietary integration

Following three examples will show how IDM benefit from their proprietary integration of manufacturing and design.

Cost reduction

The packaging cost is rising to the point could account for more than 30% of cost to produce a 130nm chip and due to the problem in heat dissipation the cost will rise even more as the process evolved. Because the new chip normally includes more functionality and runs in higher frequency, bringing out the pin is becoming a big issues. First, current package may not be able to accommodate that many gold wires. Second, even those wire can be brought out, the inductor effect among those wires will serious impact chip's performance [14]. Flip-Chip is one of the best solutions to date. IDM companies like Intel has been using Flip-Chip to resolve large pin-out for years through its in-house patented packaging technology. Flip-Chip packaging actually reduces Intel's cost in integrating multiple functionalities.

Performance enhancement

Agere reports that two key innovations were required to boost power transistor performance and reliability. The first innovation involves the elimination of defects in chips made from ultra-thin silicon wafers. Since thicker chips tend be warmer, they don't conduct heat as well as thin chips. To solve this problem Agere created a proprietary wafer scale low cost, and high yield "die (chip) thinning" technique that results in thinner and more thermally efficient (cooler) chips that are about 30 percent shorter in length as well as 50 percent thinner than all currently available
transistor products. The second technology advancement reportedly improves the transistor's performance when amplifying wireless signals. Through the creation of high-density, low resistance electrical connections, Agere is now able to fabricate transistors featuring reduced resistance and parasitic capacitance, thereby boosting each transistor's gain and efficiency parameters [15].

New product introduction

STMicroelectronics has created a technology that allows silicon-based light emitters to match the efficiency of traditional light-emitting compound semiconductor materials such as gallium arsenide (GaAs). The ability to combine optical and electronic processing on the same chip presents enormous opportunities to develop many new types of semiconductor products, especially as the technology is compatible with existing volume production process flows and equipment. The other example is that Intel takes advantage of its SiGe process to design a single chip solution for mobile phone application.

2.2 Conclusion

The factors and facts present in this chapter illustrate that semiconductor industry is in transition now. In general, foundry and fabless business models will stay. The prosperity of IDMs will depend on the scale of economy, scale of fixed fab investment and temporary advantage in steeper learning curve. Following is the trend can be observed so far:

Due to hold-up issues, IDM is unlikely to be success as a general-purpose foundry

Fabless companies are competitors of IDMs, therefore it is unlikely that fabless will award manufacturing contract to IDM's foundry. The rare exception is Xilinx to have IBM as its foundry, but Xilinx is a FPGA vendor with no product overlay with that of IBM.

The complexity of designing the manufacturing equipment and EDA tool is proportional to the progress of Moore's Law. IDM will be difficult to design equipment and EDA tools in-house.
The market capitalization of equipment or EDA companies is too big for IDM to swallow either. IDM need to work with these companies in tuning equipment and tools, but IDM can’t prohibit these companies selling the same equipments or tools to foundries and fabless firms.

**The boundary between IDM and foundry is getting blur**

Big IDM firm, like IBM is beefing up high-end specialized foundry business. In the meantime, foundry also takes virtual-IDM route. For example, UMC and TSMC invest in many fabless companies and SIP/ASIC design service firms. Their purpose is to expand the portfolio of process IPs and mitigate the risk of lacking process innovation associated with chip design [16]. By the same token, foundry also team up with traditional IDM like Motorola and AMD in developing process technology through working with IDM chip design team. In return, IDMs normally will be guaranteed to have adequate foundry capacity.

**IDM will continue to enjoy temporary advantage results from integration**

Down below 100nm the design issues that caused delays at larger geometries become less trivial. Together with the thermal issue and introduction of copper, the yield rate becomes a nightmare for foundry [10]. This new process and material require fine tuning design tools and manufacturing equipment, which will only be possible if the manufacturer work closely with chip design team. The learning curve may cause foundry continue to lag behind IDM in process improvement unless foundry came out a way to reduce transaction cost. This transaction is mainly induced by the separation of chip design and manufacturing as we mention above.

**It is a game of racing down the learning curve (i.e. ramping up the yield rate)**

It is not clear if IDM can reap all profit from the longer learning curve starting at 130nm process technology. As we mentioned earlier in this chapter that semiconductor is such a complex device, every complementary technology has to be ready in the same time for production. After the complementary technology is ready, every semiconductor manufacturer still has to work on
the most important factor, the yield rate. Figure 2.6 shows how semiconductor industry S-curve and learning curve evolve with Moore’s Law.

FIGURE 2.6

IDM may enjoy the temporary advantage in racing down the learning curve right now, but over the time, foundries will ramp up their yield rate as well. During this transition, equipment vendors will be happy to teach foundry on whatever they learn from working with IDM, since after IDM, foundries will be their only customers left. New business is always sweeter than the old one.

For those IDM can’t continue achieve process innovation or invest in new fab, they will have to migrate to fabless or fablite models (outsourcing part of demand to foundry, but keep advance process in-house).
3 The State of SoC Market

SoC (System-On-Chip) is still an evolving product segment within semiconductor industry, but was estimated to have US$34.7 billions in sales by 2007 and with compound annual growth rate (CAGA) of 20% [17]. As the name implies, it could integrates a complete system (e.g., CPU, memory and mixed signal etc…) into a single piece of silicon and shares the same manufacturing and design technologies with other high-volume customized product segments, especially ASIC product. This chapter will introduce the benefit of SoC and analyze SoC industry in the context of Porter’s Five-Force model [18].

SoC can be treated as a sub-category of ASIC. Its principal goals are to reduce design cost and enhance system integration. The primary difference between SoC and ASIC is that SoC design try to maximize reuse of existing blocks or “cores”, therefore it can minimize the amount of chip that is newly created. In many cases, SoC designers take advantage of proven working cores (in the form of standard cell) licensed from SIP vendors to accelerate the development cycle. Because of the advantage of SoC in product development cycle, functionality and cost structure, the ASIC market is converging on SoC with the exception of product segments required very high performance. At the current stage, SoC is making the inroad into applications previous only occupied by ASIC. Almost every fabless company now claims that they are either have or working on some kind of SoC product.

3.1 The benefit of SoC

With the ever advanced process technologies, more and more previous discrete components can be squeezed into a single piece of silicon. The quantitative growth in transistor density periodically causes qualitative changes in the nature of integrated circuits. As the result, not only is the price of end user system falling, but also the size of the system is getting smaller, which
indirectly push down the cost and expand SoC’s application. In addition, the quality gap between full custom and ASIC/SoC is diminishing. The industry has evolved through SSI, MSI, LSI, VLSI, ULSI and now to SoCs.

Clearly, systems-on-chips promise huge cost savings that open up new applications to mass markets. The leverage of advance process technology allows chips to be made in huge quantities at very low costs. Since most of SoC designs include one or multiple CPU/processor, SoCs are programmable. They act as platforms that enable entire product categories, not just a single product.

SoCs allow mass customization of products. The complexity inherent in a SoC, though it presents substantial design challenges, opens up new opportunities for adapting the product to the customers' needs. By changing the software on a SoC, the platform can be customized for new members of a product family or even for an individual customer.

SoCs are also the key to low-power systems. Low-power operation is essential for the battery-powered devices that are now so popular. Power consumption is increasingly important in wall-powered systems, as high power consumption and heat dissipation increase both acquisition and installation costs. Integrating multiple functions onto a single chip eliminates chip-to-chip communication that consumes precious power.

Standard CPUs and DSPs are not SoCs because they don’t solve a system problem. CPUs and DSPs are simply components in larger systems—they need memory, I/O, and multiprocessing to create full-fledged systems. As Moore’s Law drives up levels of integration, it open up new opportunities for creating custom single-chip systems; those systems use today’s components/IPs as building blocks. Figure 3.1 shows how Moore’s Law makes it possible to squeeze a number of ADSL chipset products into an integrated SoC solution over time.
3.2 New entrant and entry barrier

In SoC market, most of new entrants are fabless startup. These startups entering the market with the hope that they can develop a “killer” chip, which can sweep the market and gain the first mover advantage in the targeted application. Their view of semiconductor business is somewhat analogical to the “hit” based film making business with much less required initial capital investment. To a new fabless startup, the fixed cost in developing a SoC presents the single biggest initial investment. Since the nature of semiconductor business is high scale of economics and low marginal cost, first mover advantage is indeed significant. Adding to this thinking is the separation of fab investment and chip design cost. Without the need to reinvest in manufacturing
equipments (foundry’s responsibility), the only thing fabless need to do is to concentrate on their value add, designing and marketing a killer chip. Marketing cost is not significant to a new fabless. Since customers are equipment manufacturers and new fabless normally carries few products, there is no need for mass media advertisement or local office, at least in the beginning. The typical marketing strategy is to use local representatives/distributors and/or direct marketing.

With less marketing expense, low marginal cost and the absence of PP&E investment, it leaves the initial product development cost as the major entry barrier for new entrant. Following example will show that the entry barrier is quite low, which results in endless assault by new fabless startup and fast dropping profit margin in SoC business.

**The cost of developing a SoC**

The major cost components involved in SoC project development can be broken down into following categories.

- Labor – VLSI, hardware, software engineers and other supporting personnel.
- Tooling & materials – Photo masks and wafer slot etc...
- Design tools – EDA and emulation systems (can be reused by other projects) etc...

Of these cost components, tooling, especially the photo mask is the major piece and may require multiple sets if there is critical design error. In general, it is very difficult to have a SoC coming out of the first photo mask set and running as expected. To estimate the development cost of a SoC, we break down the development cycle showed in figure 3.2 and calculate cost as following [19].

- **Resources:**
  - VLSI Design: 60 man months
  - Hardware and Systems Design: 30 man months
  - Software Design: 40 man months
- Others: 20 man months
- Total Manpower: 150 man months

- **Material/NRE (non recurring expense):**
  
  **Case 1:** use .13u process
  
  - .13u **mask set (full) x 2:** $650,000 x 2
  - .13u wafer lot: 1 nominal E-lot - 25 wafers x $2856/wafer = $71.4K

  **Case 2:** use .18u process
  
  - .18u **mask set (full) x 2:** $225,000 x 2
  - .18u wafer lot: 1 nominal E-lot - 25 wafers x $2300/wafer = $57.5K

  **Case 1 & 2**

  Load, reference, development boards, tools, other: $96.6K

- ***Total Project Costs:***

  **Process**

<table>
<thead>
<tr>
<th>Firm</th>
<th>.18u</th>
<th>.13u</th>
</tr>
</thead>
<tbody>
<tr>
<td>Startup</td>
<td>$1,856,000****</td>
<td>$2,720,000</td>
</tr>
<tr>
<td>Fabless</td>
<td>$3,110,000</td>
<td>$3,973,000</td>
</tr>
</tbody>
</table>

- **Assumption:**

  * About $16,700/man-month for established fabless or IDM, which includes all of SG&A overhead. $8,350/man-month for new fabless startup assuming most engineers (often as part of founding team) trade 50% salary for equity. The startup scenario is also applicable to
foreign fabless firms in southeast Asia, which bears much lower labor cost structure than existing US firms.

** Average 2 full sets of photo masks.

*** Cost of licensing Silicon IP is not included due to the wide variety of IP and pricing mechanism. SIP firm may charges small initial NRE and/or per chip loyalty.

**** For a startup, the total cost might be further reduced by 30% to about $1.3 millions in .18u project, if it outsourced backend design to ASIC design service firm. Existing fabless or IDM normally won’t outsource design to 3rd party, either because this is part of their core competence, or don’t want to risk the potential schedule delay introduced by 3rd party.

![Typical SoC Development Cycle](image-url)

- System Specification
- IP Selection
- HDL & Integration
- Functional Simulation
- Synthesis & Simulation
- Place & Route
- Timing Verification
- Sign-off (tape out)
- Fab Prototype
- In System Testing
- Initial Production
- Tuned Yield Rate
- Volume Production

**Figure 3.2**
3.3 Customer Influence

Semiconductor market segments are tied to their applications and can be characterized as following.

![Pie chart showing market segments contributing to total wafer demand.]

**Figure 3.2** Source: FSA, 2001 Survey

Since SoC integrates processor core, logic components and some mixed signal etc... into a single silicon, its market mix is closely resemble to the overall semiconductor market with the exception of PC and defense sectors. Of all market segments, wired, wireless communications and consumer electronics account for more than 70% market.

**Customer influence & market segment**

Customer influence in SoC market is high. Not only is the downstream market quite concentrated, but also the big customer tends to make some of SoC for their own use. The evidence can be illustrated in following market data.

- Wired: Major customers are Cisco, Lucent, Nortel, Motorola and 3COM etc... Major applications are enterprise connectivity and telecom access (ADSL, cable modem and analog modem) etc...
• Wireless: Major customers are Nokia, Cisco, Ericsson and Samsung etc... Major applications are mobile phone, base station and 802.11 wireless LAN etc...

• Consumer electronics: Major customers are Sony, Panasonic, Philips and LG etc...

In every market segment, top three SoC customers own a disproportional market share. For example, Cisco alone owns more than 50% enterprise communication market. Motorola, Thomson and Pace together own more than 60% cable modem market [20]. Consumer electronics market looks more favorable to SoC firms, since the market is more fragmented. Unfortunately, almost every major consumer electronics vendor uses its own in-house SoC. In many cases, SoC customers are themselves SoC suppliers e.g. Motorola and Philips.

**Low switching cost**

Quite often, it is not difficult for customers to switch SoC suppliers. A specific SoC can normally be replaced by a combination of another SoC and discrete components with some software changes. This presents low switching cost to customers. Some market segment does have higher switching cost, e.g. broadband communications application. In this case, customers often simultaneously design multiple products using different SoC solutions to enhance their bargaining power. The cost of designing another PCB before shipping product is negligible.

**Less product differentiation**

Because of Moore’s Law, a single silicon can now accommodate a lot more functionality now, therefore more targeted application can be support by a SoC. It is not rare to see that more than 30% of features in a SoC are not taken advantage by the end user product. This results in highly overlaid feature set and lack of product differentiation.

3.4 Competition

With more than 15 IDMs and 600 fabless firms worldwide, the market is very crowded. Still the new fabless startup keeps popping up. This is because of the versatile application of SoC. In
most of market segments, either there is a shortage of “killer application” or the application is still evolving. In other words, everyone is still figuring out what is the “dominant design”. In the context of S-curve, most of market segment is still in the ferment stage. There is going to have lots of consolidation in the near future.

3.5 Foundry’s Influence

The top three foundries excluding IBM account for more than 70% of total capacity. Foundry’s influence is enormous. Especially, due to process technology dependency, it is not easy for fabless firm to switch foundry partner. It will take at least 6 months for fabless companies to retool IP library and process dependent parameters. Figure 3.4 shows the market share of foundries.

![Primary Foundry Partners](image)

**Figure 3.4**

3.6 Substitute

An FPGA (Field Programmable Gate Array) and PLD (Programmable Logic Device) are general-purpose, multi-level programmable logic device that is customized in the package by the
end users. Comparing to SoC, the advantage of using FPGA/PLD is low tooling cost, rapid turnaround, low testing cost and cost reduction over product life cycle, whereas its disadvantage is large die size and slow speed. The disadvantage of large die size makes FPGA/PLD an infeasible solution for high-volume application.

Lately, the environment is becoming more favorable to FPGA/PLD. Due to following factors, FPGA/PLD begins to make inroad into some market segments traditionally occupied by ASIC/SoC.

**High tooling cost to design an ASIC/SoC**

For .13u, a full mask set costs $650,000. It usually requires the cost of about 2 full mask sets to complete a SoC product. The more functionality got squeezed into a chip, the higher possibility to have design mistake. The more mistakes been made, the more mask sets required.

**Short downstream product life cycle results from the shortage of dominant designs**

After the introduction of TV, VCR, PC and mobile phone (GSM version), IT related industries are still searching for the next big idea. Since most of downstream market segments are still looking for killer application, SoC design houses normally take the trial-and-error approach to develop product, otherwise they may risk losing first mover advantage. This strategy produces a wave of SoC product with little differentiation and adds huge extra development cost for every firm. With the advantage of reprogrammability and no tooling cost, FPGA/PLD is increasingly viewed as a viable solution at this ferment stage.

**New FPGA/PLD architecture**

Leading FPGA/PLD vendors have been busy in integrating more functionality into their products. The high-end FPGA/PLD now integrates with processor core, basic I/O block and more commonly used logic components, which means that number of functionality is up, but the die size stays the same or smaller. This new form of FPGA/PLD also called SoRC (System-on-a
Reprogrammable). In certain aspect, this new architecture is closer to SoC than to its traditional FPGA/PLD siblings.
4 The Impact of FPGA/PLD to SoC Industry

According to Gardner Dataquest 2001 report, close to 50% of today ASIC designs are SoC with a good portion of the ASIC design starts with in the capabilities of today’s system-level FPGAs. One-in-three system designers now use FPGAs with more than 100K gates. In addition, because the price of FPGA/PLD continues to decline, one in three of FPGA/PLD designs are used in production today.

4.1 New Development in FPGA/PLD

Because design reuse is about planning for the future, the term System-on-a Reprogrammable FPGA/PLD (SoRC) is used for SoC implemented in an FPGA/PLD. This term is used to define both full and partial system-level designs since the challenges facing these FPGA/PLD designers are almost identical. The advantage for FPGA/PLD vendor to integrate SoC with FPGA/PLD is:

- Enhance the programmability due to the addition of processor core.
- Reduce the die size since some commonly used logic components are implemented in standard cell or full custom (ASIC) and integrated into the same silicon.
- Reduce the power dissipation in FPGA/PLD.
- Increase the over all system performance (clock speed)

With the right targeted application and the integration of required logic component and processor core, SoRC could indeed get the best out of both SoC and FPGA/PLD worlds, namely, the small die size, low power consumption and high speed of SoC and fast time to market, low development cost and good programmability of FPGA/PLD. For above reasons, FPGA/PLD vendors (Xilinx and Altera etc…) all implement following product marketing strategy.

- Target applications with low-to-mid volume potential for SoRC products.
- Integrate full-custom logic components and processor core to support targeted applications.
• Utilize the most advance manufacturing process (foundry) to manufacture their products.

That way they can offset the disadvantage of bigger die size (comparing to the ASIC or SoC with equivalent functionality).

4.2 SoC or SoRC: A Question of Total Cost

From section 4.1, we know that all FPGA/PLD vendors now target applications with low-to-mid volume potential for their SoRC products, but exactly what is “low-to-mid” volume? The answer is in the comparison of the total cost between SoC and SoRC.

Total Cost (TC) = Fixed Cost (FC) + Variable Cost (VC)

Cost of SoC

• Fixed cost

This is the expenditure of labor, tooling and test/system. As it was showed in section 3.2, a SoC project using 0.18u process will cost about $3.11M for the existing fables/IDM firms.

• Variable cost

The variable cost is a function of die size and yield rate. Depends on the number of features integrated in a chip, a wafer may consist of tens or hundreds of chip. Since the density of transistor double for every generation of process technology, the number of chip with same features in a wafer will be doubled as well. For example, at present:

1. 0.18u: $2,300/wafer
2. 0.13u: $2,856/wafer
3. Number of chip on a 0.18u wafer : Number of chip on a 0.13u wafer = 1 : 2

This result in about 40% cost reduction by simply manufacturing chip using next generation process technology.

• Other cost factors
Yield rate is the result of tuning effort. It may take 3 ~ 4 months to tune the yield rate from about 50% in the first wafer to 90%. A SoC project normally has development cycle of one and a half to two years.

Time to market is another factor. When business is slow, many companies think they have the extra time they need to create products using ASIC technology, trying to get the lowest possible production cost. The rationale is that the long design cycle of an ASIC won’t hurt them because the slow economic conditions allow more time before new products need to be introduced. In fact, these very same slow economic conditions may put pressure on companies to very quickly come up with innovative products when their end markets recover. When business starts to improve, it is the companies that get their new products to market first that will reap the rewards of the upturn. Designing with ASICS, may not be able to quickly modify product to meet new market needs. Programmable logic provides the flexibility to quickly develop products that may allow company to realize the maximum profit from an improving economy.

Cost of SoRC

- Fixed cost

FPGA development required much less engineering effort and resources than that of SoC.

Figure 4.1 show the difference in design flow.
Comparing to SoC development, FPGA commands following cost saving:

1. No repeated structure verification

2. No fab prototyping. FPGA do not require long lead times for prototypes or production parts - the FPGA are already on a distributor’s shelf and ready for shipment

3. No yield rate tuning. The reason why yield rate tuning does not contribute to fixed cost is that the same FPGA is used by all of different customers/projects. FPGA vendors will shoulder the cost. Since FPGA vendors aggregate demand from all customers and spread development cost to every chip they ship, the development cost is insignificant
comparing to volume. For example, since its introduction in 1998, Xilinx has shipped more than 50 million Spartan Series devices.

4. Rapid design turnaround. FPGA offer customers much more flexibility during the design cycle because design iterations are simply a matter of changing the programming file, and the results of design changes can be seen immediately in working parts.

5. Reusable design and system test tools. In general, the major cost in developing a FPGA is labor expense. The time and labor required to complete a FPGA project is roughly 1/3 comparing to that of SoC [21]. That translates into about $835,000. Since a SoRC chip includes both SoC and FPGA, only FPGA portion need to be redesigned. The SoC functionality embedded in a SoRC is pre-designed and fabricated by the vendors. There is no need for additional engineering effort.

In conclusion, the only major fixed cost in developing a SoRC is the labor cost related to VLSI engineering. From the example in section 3.2, this cost will be less than $835,000 and is only proportional to the portion of SoRC functionality implemented in the FPGA.

- Variable cost

The variable cost of a SoRC is only related to its die size. Yield rate is always high –more than 90%, since that is vendor’s job and best interest to get the high yield rate.

**Cost Comparison: SoC vs. SoRC**

To conduct the cost comparison between SoC and SoRC, there is certain assumption has to be made.

1. Comparing to fabless SoC firms, FPGA vendors are normally at least one generation ahead of process technology. For example, both Xilinx and Altera are manufacturing their high volume product in 0.13u process, while most of fabless firms and IDMs are
still in the process to migrate their products from 0.25\(\mu\) to 0.18\(\mu\) process. Xilinx even announced the availability of its Sprtan-3 FPGA in 0.09\(\mu\) and 12-inch wafer process.

2. Both SoC and SoRC use standard cell to implement features. Due to the flourish of SIP industry, there is ample features can be licensed in the form of standard cell. Most of SoC designers do use SIP. Take the home networking SoC showed in figure 4.2 as an example. In this design, processor core, the ARM940T is licensed from ARM in the form of full custom ASIC block, which occupies very small die area. RAM and ROM are also full custom ASIC block. Rest of most functional blocks is implemented in the form of standard cell licensed from SIP vendors.
Here we assume all functionality in SoC other than processor core, RAM and ROM are implemented in standard cell. Assume that there are following three alternatives to implement a SoRC with mix of standard cell blocks and FPGA.

1. All standard cell(functionality) in the SoC was implemented in SoRC using FPGA
2. Half of standard cell implemented in FPGA, remaining half kept in standard cell
3. 25% of standard cell implemented in FPGA, remaining 75% kept in standard cell
Figure 4.3 illustrates the cost comparison in terms of die size between SoC and the equivalent SoRC implementation.

<table>
<thead>
<tr>
<th>% of standard cell implemented in FPGA</th>
<th>SoC die size</th>
<th>SoRC die size</th>
<th>Die size comparison SoC : SoRC</th>
<th>Variable cost comparison SoC : SoRC</th>
<th>*Fixed cost of SoRC development</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>1 0.18u</td>
<td>10 0.18u</td>
<td>1 : 10</td>
<td>1 : 10</td>
<td>$835,000</td>
</tr>
<tr>
<td>50%</td>
<td>Standard cell 0.18u</td>
<td>5.5 0.18u</td>
<td>1 : 5.5</td>
<td>1 : 5.5</td>
<td>$417,500</td>
</tr>
<tr>
<td>50%</td>
<td>Standard cell 0.18u</td>
<td>2.75 0.13u</td>
<td>1 : 2.75</td>
<td>1 : 3.3***</td>
<td>$417,500</td>
</tr>
<tr>
<td>25%</td>
<td>Standard cell 0.18u</td>
<td>3.25 0.18u</td>
<td>1 : 3.25</td>
<td>1 : 3.25</td>
<td>$208,750</td>
</tr>
<tr>
<td>25%</td>
<td>Standard cell 0.18u</td>
<td>1.625 0.13u</td>
<td>1 : 1.625</td>
<td>1 : 1.95***</td>
<td>$208,750</td>
</tr>
</tbody>
</table>

Figure 4.3
* As mentioned earlier in this section, the time and labor required to complete a FPGA project is roughly 1/3 comparing to that of SoC. That translates into about $835,000 and is proportional to the portion of functionality implemented using FPGA.

** A feature implemented in FPGA will required 10 times bigger in die size than that of does in standard cell [22].

*** Cost reduction from migrating to next generation process technology is about 40%.

**Minimum Efficient Scale (MES)**

Following formula is derived to find out what is the minimum lifetime volume required to justify a SoC development project.

SoC variable cost + SoC fixed cost = SoRC variable cost + SoRC fixed cost

- **Case 1**

  **Assumption:**

  - SoC, 0.18u
  - SoRC, 0.13u, 50% of functionality implemented in FPGA
  - The die cost of above home networking SoC is about $3. In general, the die cost of a SoC can be derived as following:

    \[
    \text{Die size} = (\text{sum of gate count of all standard cell}) \times \text{transistor count/gate} \times \text{area size/transistor}
    \]

    \[
    \text{Die cost} = \text{wafer cost} / (\text{yield rate} \times \text{number of die per wafer})
    \]

  **The result**

  \[
  \text{SoC die cost} \times \text{volume} + 3,110,000 = \text{SoC die cost} \times (\text{SoRC die cost} / \text{SoC die cost}) \times \text{volume} + 417,500
  \]

  \[
  \Rightarrow 3 \times (3.3 - 1) \times \text{volume} = 3,110,000 - 417,500
  \]

  \[
  \Rightarrow \text{volume} = 390,217 \text{ units}
  \]
Conclusion

If the volume is less than 390,217 units, it is cheaper to design the chip in SoRC.

- Case 2

Assumption:
- SoC, 0.18u
- SoRC, 0.13u, 25% of functionality implemented in FPGA
- SoC die cost = $3

The result

SoC die cost * volume + 3,110,000 = SoC die cost * (SoRC die cost / SoC die cost) * volume + 208,750

=> 3 * (1.95 -1) * volume = 3,110,000 - 208,750

=> volume = 1,017,982 units

Conclusion

If the volume is less than 1,017,982 units, it is cheaper to design the chip in SoRC.
The timing factors

- **Time to market**

The importance of time to market can be summarized in two factors, "market share" and "profit margin". The Urban et. al. (1986) [23] study shows the relationship between market share and order of entry in Table 4.1. In another study [24] conducted by Reinertsen (1983), it shows that in the high-technology market, a six-month delay in product delivery cut the lifetime profits of a product by 33%.

### Market Shares and Order of Entry

<table>
<thead>
<tr>
<th>Entry Order =&gt;</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Brand</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>59%</td>
<td>41%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three</td>
<td>44%</td>
<td>31%</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td>Four</td>
<td>36%</td>
<td>25%</td>
<td>21%</td>
<td>18%</td>
</tr>
</tbody>
</table>

Table 4.1
• Product life cycle and Moore’s Law

Moore’s Law has great impact on the life cycle of semiconductor products. It effectively reduces the product life cycle of most SoC to about two years. In other words, if a leading SoC product does not have a cost reduced and performance enhanced successor in 18 months, competition will surely introduce a similar product with lower cost (40% down) and higher performance as a result of Moore’s Law. Although it doesn’t mean that a company can not sell a same product for more that 18 months, it is in the best interest of the company to plan a follow-up product using next generation process technology. The reason is higher profit margin (assuming customer still willing to pay the same price) and deterring new entrant. In considering above timing factors, we re-calculate case 2 as following.

• Case 2 + timing factors

Assumption:
- SoC, 0.18u
- SoRC, 0.13u, 25% of functionality implemented in FPGA
- Original SoC die cost = $3
- Typical SoC profit margin is 60%, (price-cost)/price, at its introduction.

The result

The introduction of a SoC product will be at least 6-month behind the equivalent SoRC, therefore the lifetime profit of the SoC product will be cut by at least 33%. This will be the equivalent of 50% jump in cost.
- No delay: (price-cost)/price = 60% => cost = 0.4 * price
- Delay > 6 months: (price-new cost)/price = 60% * (1-33%) => new cost = 0.6 * price
- Cost increase: (0.6 – 0.4)/0.4 = 50%
New SoC die cost * volume + 3,110,000 = Original SoC die cost * (SoRC die cost / Original SoC die cost) * volume + 208,750

=> 3 * 1.5 * volume + 3,110,000 = 3 * 1.95 * volume + 208,750

=> volume = 2,149,074 units

Conclusion

If the volume is less than 2,149,074 units, it is cheaper to design the chip in SoRC. In the worst case scenario, the product life cycle might tie to the of Moore's Law, i.e. 18 months. That means the company has to sell more than 2,149,074 units within 18 months to justify the SoC development, otherwise SoRC is a better choice.

Note: With the latest FPGA/PLD architecture, the ratio of die size required to implement a FPGA with equivalent functionality of a SoC is about 5:1. This further reduces the cost of FPGA/PLD devices.
4.3 Conclusion

Since their introduction in 1984, FPGA have become one of the most popular implementation media for digital circuits and have grown into a $2.5 billion per year industry. As process geometries have shrunk with Moore’s Law, the logic capacity of FPGA has greatly increased, making FPGA a viable implementation alternative for larger and larger designs. Due to the new architecture of integrating FPGA, processor and core logic into a single silicon, SoRC is making the inroad into high volume applications, which traditionally occupied by SoC. Through the discussion in this chapter, the competition between FPGA/PLD and SoC industries can be inferred as following.

- Due to the accelerated downstream product life cycle, SoRC will begin to take market share away from SoC in the mid-to-high volume market segment. Traditional FPGA/PLD will continue dominate the low volume market.

- Higher fixed cost in designing a new chip will reduce the incentive to implement SoC. This will further help FPGA/PLD to expand market share

- The new reality of Moore’s Law —longer duration to double transistor density— will hurt FPGA/PLD vendors more than SoC vendors. This is because, first, the fixed cost in developing a SoC can now be spread into longer product life cycle. Second, FPGA/PLD die size can’t be squeezed fast enough to compensate the weakness of bigger die size and lower performance. The prolong technology transition (e.g. from 0.13u to 0.09u) is a double edge sword. On one hand, it forces SoC vendors to stay with less advanced process, otherwise will be hit by high tooling cost. On the other hand, it takes longer for FPGA/PLD vendors to reduce the die size i.e. slow down the pace of cost reduction. As a result, the profit margin of entire semiconductor value chain will be lower. The free ride of productivity gain from Moore’s Law will slow down.
• IDMs could enjoy temporary advantage (steeper learning curve) during the transition of process technology. It will likely squeeze fabless SoC vendors in high volume market segment. Ex. Intel in mobile phone (use SiGe process), TI in DSP and xDSL, Motorola in general-purpose microprocessor markets. Due to the ownership of fab, IDM also has advantage in high volume and long life-cycle product segment. Figure 4.4 shows current market segmentation.

<table>
<thead>
<tr>
<th>Product Life Cycle</th>
<th>Long</th>
<th>Short</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>IDMs</td>
<td>IDMs &amp; Fabless</td>
</tr>
<tr>
<td></td>
<td>ASIC &amp; SoC</td>
<td>SoC &amp; SoRC</td>
</tr>
<tr>
<td>Low</td>
<td>Fabless</td>
<td>Fabless</td>
</tr>
<tr>
<td></td>
<td>FPGA/PLD</td>
<td>FPGA/PLD</td>
</tr>
</tbody>
</table>

*Figure 4.4*

• Due to organizational rigidity, IDMs may not be able to compete effectively in the high-volume but short life-cycle product segment. If the product life cycle is as short as two years or less and market size is less than three millions, SoRC will have the edge over SoC.
5 Summary

This thesis presents an overview of the semiconductor industry. It presents a snap shot on the current industry dynamics, especially the relationship among IDM, foundry and fabless firms. It goes on to discuss the ever growing SoC market segment and its impact on other semiconductor products. Lastly, it evaluates the feasibility of using SoRC as an alternative solution of SoC.

Contrast to the conventional wisdom regarding to hi-tech industry, semiconductor is becoming an industry competing on cost. The fixed and variable costs are the central topics of semiconductor industry and the driving force is Moore’s Law.

To build a new fab, foundries or IDMs have to justify the huge fixed initial investment. The required economics of scale in this industry is enormous. Due to the low variable cost in producing additional chip, the product development expense becomes one of the biggest cost component for fables companies. Adding into the fuel is the Moore’s Law. Not only does it push IDM and foundry in building new fab, but also has profound effect on the semiconductor product life cycle and its minimum efficient scale (MES). Together Moore’s Law and learning curve (yield rate) dictate the fixed and variable cost of a semiconductor product and further influence companies’ product development decision.
References

[14] Intel company announcement, Nov. 28, 2002
[19] Conexant Personnel Computing Division, project management estimate, 2002
[21] Interview with Cisco ASIC design project manager, Benjamin Tandiono, April 2003