GILDER
TECHNOLOGY
REPORT

November 2003 / Vol. IX No. 11

Semiconductor Investment and the Future

The semiconductor business is now on the verge of brisk expansion, with enormous growth potential in China and in wireless access and consumer devices

Inside:

Investing is a hot topic. We read a lot about it and we hear a lot about it. Stocks, bonds, commodities, futures, currencies, gold, mutual funds, index funds, hedge funds, large-caps, small-caps, puts, calls, shorts, longs, market sectors, etc. There's Fed policy, interest rates, inflation, deflation, taxes, tax credits, capital gains, dividends, IRS, SEC, GDP growth, CPI, business forecasts, ITC, WTO, globalization, and thousands of analysts with opinions for sale. What a confusion of things to think about! Here's how I simplify the financial zoo.

I am immersed in semiconductors—it is an industry like no other. Semiconductors invade *other* markets. Think about that. Soft drinks, fast foods, carpets, furniture, tractors, office supplies, pork bellies, and automobiles don't do that—can't do that. If you invest in Coca-Cola, you are betting that its executives will displace Pepsi, will get into refrigerators they have never been in, will convert tea drinkers to soft drink addicts, and will reduce costs in manufacturing and in delivery. The same is true of the carpet business—there can be new floors, better fabrics, better production processes. Coke can spill over into the carpet business, but not in a way that benefits the soft-drink industry (or carpets). That's how the semiconductor industry differs. Semiconductors have invaded automobiles, consumer appliances, games, machine tools, electric motors, toys—just about anything you can name.

Steel once dominated the cost of an automobile. Today, the automotive manufacturer's investment in electronic content exceeds the cost of the car's steel. The electronic portion of the invaded segment can grow rapidly even in a mature industry. Automotive electronics grows faster than the automobile industry.

The semiconductor industry suffers bust cycles that drive away investors. But its cumulative growth for the last forty years, including 2001's downturn of almost 40 percent, is better than 12 percent per year. That means the average company in the semiconductor sector doubles every six years. By contrast, worldwide automobile production grew by 3.5 percent per year during the same period (doubling only every twenty years).

So, with this as background, here's an investment strategy that doesn't require learning and tracking all that CPA-stuff in the first paragraph. Invest in the stock market. Invest for the long term. Invest in technology companies. A random selection of technology stocks, kept for ten years, should appreciate at 12 percent. Spend your time doing something interesting instead of working put-and-call spreadsheets.

So far so good, but you can do better than 12 percent. Learn something about where the semiconductor industry is headed, for example, by reading the *Gilder Technology Report.* Learning about technology helps you separate mature and moribund companies from growth companies. The low or negative growth of mature and moribund companies weighs the industry average *down* to 12 percent. Buy a broad

CONTINUED ON PAGE 2

selection of small-cap technology stocks and hold them for the long term. You still have to do homework on technology, on corporate leadership, and on financial fundamentals. But the problem is interesting and tractable. Some companies will disappoint, but some will grow at 30 percent per year. Since we cannot know the future, there's value in diversification.

By all indicators, the semiconductor business is now on the verge of brisk expansion, with enormous growth in emerging economies such as China and in wireless access and untethered consumer devices. None of today's devices is good enough; they'll all be replaced with better ones in the next few years. In addition, digital is displacing analog in media applications. The entire photo, film, and video industry is moving from analog and wet-chemical processing to digital. New equipment, more digital storage, more networks, more computers, more software. Electronic tags (RFID) will first replace bar codes in the supply chain and will then move into a broad range of applications. Microelectromechanical sensors and actuators will integrate information systems with the physical world to bring benefits we cannot imagine today.

Biological systems do amazingly complex tasks operating at room temperature on tiny energy budgets. Forty years of integrated-circuit scaling have brought our clunky electronic systems close to the molecular granularity of biological systems. Soon, engineers will be co-opting solutions from nature. We'll have generations of efficient electromechanical devices on the way.

Semiconductor's three-event track meet

The state of the semiconductor business is like a three-event track meet. The events are the transistor marathon, the logic mile, and the memory hurdles.

The transistor marathon

What's happening

The transistor event is a marathon with a single entrant, the transistor. The industry has been brute-force shrinking the transistor for forty years. We're near the finish line. *Transistors are now good enough for most applications.* Make them too large and they use too much energy; make them too small and they use too much energy (leak too much).

Why it's happening

For decades, transistors weren't good enough. Transistors were big and slow and they burned a lot of energy switching on and off. Smaller transistors were faster. More transistors fit on a chip, so, while an individual transistor's switching energy went down, the chip's overall power use went up. Rising chip power was OK because most electronic systems got their power from the wall.

Four developments—foundries, emerging countries, untethered systems, and rising chip costs—combine to create what I call the "value" transistor, good enough for the job and optimized not for small feature size but for cost-performance-per-watt.

Foundries. Foundry production of chips is overtaking production by integrated device manufacturers (IDMs). The difference between foundries and IDMs is that foundry production is demand driven—foundries build to meet customer demand—while IDMs drive production following internal business models. At a foundry, customer orders drive the mix of semiconductor processes (e.g., the percent of wafer starts at 350-nm, 250-nm, 180-nm, 130-nm, and 90-nm transistor sizes). At an IDM, such as **Intel** (INTC), **Sony** (SNE), or **Micron** (MU), corporate fiat drives the mix of transistor sizes. In the heyday of IDMs, owning a fab conferred a competitive advantage through shrinking transistors.

Growing markets in emerging countries. Today the market is splitting into tethered and untethered systems. The fastest growing markets in tethered systems are in emerging countries where huge populations are buying their first appliances. Tens of millions of refrigerators, blenders, electric toothbrushes, microwave ovens, and hair dryers. Consumer appliances don't need 90-nm transistors, they're happy with the cheapest transistors they can get. Today, the cheap transistors come from old, fully depreciated manufacturing processes, such as 250 nm or 180 nm.

Untethered systems. For untethered systems, smaller transistors aren't always better. As the transistor shrinks, its active power (the energy the transistor uses in switching on and off) decreases, but it leaks more, even when it's in its off state. At 90 nm (the leading-edge in 2003) the transistor's active power is about 10,000 times its leakage power (operating at 1 GHz). In two generations, at 45 nm, the active power will decrease and the leakage will increase so that the active power will be 150 times the leakage power. These numbers make it sound like leakage power isn't a problem, but consider that the number of transistors on a high-end chip is in the hundreds of millions. If a majority of the transistors are busy, then the active power matters. If few of the chip's transistors are busy and most of the transistors are sitting around leaking, then the leakage matters. In today's biggest chips, most of the transistors are idle—essentially all of them are leaking. If one percent of the chip's transistors are busy ten percent of the time, then, at 90 nm, the active power is 10 times leakage power. At 45 nm, the leakage power becomes 7 times the active power.

Among applications, there's a huge spread between how many transistors are active and how many are leaking. Web servers and Internet routers have more active transistors and fewer leaking transistors. Untethered systems, such as smoke detectors and cell phones, have more leaking transistors and fewer active transistors. The cell phone's job, like the classic **Maytag** (MYG) repairman, is mostly to sit idle. While the argument might be over whether the point is reached at 130 nm, at 90 nm, or at 65 nm, the shrinking transistor eventually passes a point where its leakage power overtakes its active power. Beyond that point, smaller transistors are worse for that application. As more transistor sizes become available, more applications have the right transistor available. They don't need and won't pay for smaller transistors.

Rising chip costs. Chip costs divide into fixed costs and variable costs. Including the manufacturing plant, its equipment, semiconductor-process development, chip design, and masks, fixed costs are the same whether you produce one chip or a billion. While plant, equipment, and process-development costs are amortized across all the wafers that the plant makes, chip-design cost and mask cost are amortized across the wafers for one particular chip design. The costs for processing one wafer through the plant, variable costs, are rising slowly, almost independent of the process generation. Fixed costs approximately double with each process generation. For the first thirty years of semiconductor manufacturing, in spite of their doubling, fixed costs remained small relative to variable costs. But for some applications, fixed costs now contribute more to total cost than variable costs do.

Tough times for leading-edge suppliers may mean boom times for second-tier suppliers, such as Semitool and Ultratech

Higher fixed costs need larger markets. As a back-of-the-envelope example, suppose you want the mask cost to contribute no more than ten cents to the product cost. If the mask cost is $18,000, then your market will have to be at least 180,000 chips. If mask cost is $2,000,000, your market will have to be at least 20,000,000 chips.

Back when masks cost $18 thousand, the market didn't have to be large to return a profit. Today's masks cost about a $1 million. When masks pass $2 million, some applications are left behind—unable to afford the more advanced processes. As applications are left behind, fewer applications remain to pay the costs of new process generations.

Today, already-available "value" transistors are good enough for the rapidly growing appliance markets in emerging economies, and smaller transistors may be worse for untethered applications, such as cell phones, that have huge numbers of idle leaking transistors and few busy transistors. The larger minimum markets, implied by the higher costs to make smaller transistors, make smaller transistors unaffordable for more and more applications.

Consequences

Integrated device manufacturers, such as Intel and **Texas Instruments** (TXN), will follow their shrinking-transistor business models beyond the needs of their customers. The IDMs' continued investment in shrinking transistors blinds them to the opportunity for innovation in other areas, such as three-dimensional circuits and wafer stacking. These opportunities are most likely to be exploited by startups such as **Matrix Semiconductor**, **Tezzaron Semiconductor**, and **Ziptronix**. Foundries, such as **Chartered** (CHRT), **Taiwan Semiconductor** (TSM), and **United Micro Electronics** (UMC), will increasingly dominate chip production, but with a slowly changing mix of semiconductor processes.

Foundries in emerging economies, such as **Grace Semiconductor**, **GMSC**, and **SMIC** in China, will grow rapidly. The fastest growth in foundry capacity will be in trailing-edge semiconductor processes, for consumer markets in emerging economies. In this, Taiwan Semiconductor and UMC will be hampered by Taiwan's export restrictions, which tie the companies' foreign investment in trailing-edge capacity to their domestic investment in leading-edge capacity.

New-equipment-order backlogs of the semiconductor-equipment makers, such as **Applied Materials** (AMAT), **Lam Research** (LRCX), **KLA Tencor** (KLAC), and **Novellus** (NVLS), were once a leading indicator of industry health. Not so in the future. The market for high-margin, leading-edge processing equipment will decline as the market for refurbished equipment and for trailing-edge equipment grows. Tough times for leading-edge suppliers may mean boom times for second-tier suppliers, such as **Semitool** (SMTL) and **Ultratech** (UTEK).

The end of the shrinking-transistor marathon isn't the end of the industry. It only means that the right transistor is available for most applications. Innovation will come from better use of transistors rather than from shrinking them. There's a lot of gloom and doom in the industry over the anticipated end of Moore's law. Moore's-law progress will slow as more applications find their value transistor. That's a good thing, as it lowers the costs of innovation. Gloom-and-doom predictions come from people seeking funding for quantum computing, molecular computing, optical computing, nano-computing, and other research projects.

The logic mile

What's happening

The logic event has the greatest diversity of entrants: application-specific integrated circuits (ASICs), application-specific standard products (ASSPs), microprocessors, digital signal processors (DSPs), and programmable logic devices (PLDs). ASICs claimed an early lead in high-volume, cost-sensitive, performance-oriented systems, such as game consoles, MP3 players, GPS (global positioning system) receivers, and cellular phones. ASSPs gained ground on ASICs as custom chip development became more expensive because ASSPs share their development cost among several system makers.

Microprocessors set the pace for the bulk of consumer applications-—systems that are cost-sensitive but that are not performance-oriented: microwave ovens, toasters, irons, watches, and smoke detectors. DSPs, the fastest growing segment of the microprocessor market, augment applications that need more performance than a microprocessor can deliver, but don't need custom logic. Microprocessor-based implementations are so entrenched in the design community that software programming and problem solving have become

TELECOSM TECHNOLOGIES

Altera	(ALTR)
Analog Devices	(ADI)
ARM Limited	(ARMHY)
Avanex	(AVNX)
Broadcom	(BRCM)
Cepheid	(CPHD)
Chartered Semiconductor	(CHRT)
Ciena	(CIEN)
Corvis	(CORV)
Cypress	(CY)
Energy Conversion Devices	(ENER)
Equinix	(EQIX)
Essex	(EYW)
EZchip	(LNOP)
Flextronics	(FLEX)
Intel	(INTC)
JDS Uniphase	(JDSU)
Legend Group Limited	(LGHLY.PK)
McDATA	(MCDTA)
Microvision	(MVIS)
National Semiconductor	(NSM)
Qualcomm	(QCOM)
Samsung	(05930.KS)
Sonic Innovations	(SNCI)
Sprint PCS	(PCS)
Synaptics	(SYNA)
Taiwan Semiconductor	(TSM)
Terayon	(TERN)
Transmeta	(TMTA)
United Microelectronics	(UMC)
VIA Technologies	(2388.TW)
Wind River Systems	(WIND)
Xilinx	(XLNX)

Note: The Telecosm Technologies list featured in the *Gilder Technology Report* is not a model portfolio. It is a list of technologies that lead in their respective application. Companies appear on this list based on technical leadership, without consideration of current share price or investment timing. The presence of a company on the list is not a recommendation to buy shares at the current price. George Gilder and *Gilder Technology Report* staff may hold positions in some or all of the stocks listed.

Altera (ALTR)

PROGRAMMABLE LOGIC DEVICES

NOVEMBER 17: 23.10, 52-WEEK RANGE: 10.30 – 23.44, MARKET CAP: 8.75B

In product news:

• Korea's LG Electronics, a major provider of CDMA mobile phones and infrastructure, selected Altera's Stratix HardCopy devices for its 3G basestations. "HardCopy" refers to Altera's ASIC alternative. Engineers first use an Altera Stratix field programmable gate array (FPGA) to design and test the system, then cement the design in the top two metal mask layers of the HardCopy chip. HardCopy fits into the system just like the FPGA but is actually 50% faster. Time to market is much faster than application specific integrated circuits (ASICs), and silicon errors are all but eliminated.

• Micron and Altera announced a new interface for 400-Mbps double data rate (DDR400) DRAM and FPGAs.

• Innocor and Altera announced a major new serial protocol, dubbed SerialLite. "Intended for systems used in wireless and wired communications, computer, video broadcast, data storage, industrial, and automation applications," the companies noted, "SerialLite offers a rich, scalable feature set including scalable data transfer rates extending into the hundreds of gigabits per second (Gbps)." SerialLite will be an open and free standard.

Analog Devices (ADI)

RF ANALOG DEVICES, MEMS, DSPs

NOVEMBER 17: 45.24, 52-WEEK RANGE: 22.58 – 48.75, MARKET CAP: 16.62B

The company introduced a new analog-to-digital converter (ADC) for 3G wireless basestations, satisfying all three standards (cdma2000, WCDMA, and TD-SCDMA). The ADC can sample a 200 MHz signal with 14-bit resolution, yielding 92 megasamples per second (MSPS). Analog has also integrated a digital down converter (DDC) on the same chip, reducing component costs by more than 20%.

Analog reports earnings on November 18.

Broadcom (BRCM)

BROADBAND INTEGRATED CIRCUITS

NOVEMBER 17: 35.95, 52-WEEK RANGE: 11.86 – 37.65, MARKET CAP: 10.89B

BRCM stock is up more than 200% in the last year but still has a price-to-sales multiple in line with its fabless semiconductor peers. The company is still strategically positioned in key broadband markets and in the emerging market for "Entertainment Servers" (see GTR, October 2003) and home networking.

In product news:

• Broadcom announced its "54g" WiFi chip has been incorporated into three new wireless LAN products from Microsoft. "54g" is the moniker for Broadcom's 802.11g product that boosts top WiFi data rates to 54 Mbps from the 11 Mbps offered by the original 802.11b standard. "54g" also operates in the 2.4 GHz frequency band and is backward compatible with the installed base of 802.11b products.

• The company is shipping its new asymmetric digital subscriber line (ADSL) chip to China, where the DSL market will double in 2003.

• The company also shipped its 800 millionth Ethernet port.

EZchip (LNOP)

10 GIGABIT NETWORK PROCESSORS

NOVEMBER 17: 8.89, 52-WEEK RANGE: 3.88 – 11.20, MARKET CAP: 64.82M

We noted some good news last month, but the wider investing world seemed to miss it, so we will say it again: wireless giant Nokia announced it will use EZchips in its new 3G wireless aggregation platforms. Two weeks after EZchip first announced the deal, Dow Jones published the story. The stock spiked 20% but has now retreated.

For three years we have marked EZchip as one of the most promising young chip companies in the world. Now, with volume shipments starting in the next month or so, the moment of truth arrives, and we will be able better to gauge the company's acceptance in the market. Until now, revenues from sample chips and software tools have been negligible. In a November 4 Reuters article, CEO Eli Fruchter seemed to say by the end of the year he expects revenues to reach "several million dollars per quarter." But the article was later corrected to read that revenues should reach several million per quarter *by year end* 2004, not 2003. Fruchter tells us, "Things look very good for us but revenue ramp is going to take some time."

Intel (INTC)

MICROPROCESSORS, SINGLE-CHIP SYSTEMS

NOVEMBER 17: 32.23, 52-WEEK RANGE: 14.88 – 34.51, MARKET CAP: 210.53B

The company announced what it calls a major advance in materials and manufacturing capable of sustaining Moore's law into the next decade. We dub it "Special K." As Nick Tredennick writes this month, today's shrinking chip transistors are leaking too much current, wasting energy and battery life. Remember the basics: A gate turns the transistor on and off. The gate dielectric is an insulator that controls the flow of current under the gate. The problem is that silicon dioxide, the material used to insulate gates for the last 30 years, no longer insulates as its thickness is decreased. Electrons flow right through the walls (which have reached just 5 atoms thick in research settings). Intel's innovation is the discovery of a new, but as yet unspecified, "high k" dielectric insulator that is 100 times more resistant to leakage than silicon dioxide. Because today's polysilicon gate structure is incompatible with Intel's new "high k"

insulator, however, the company also had to invent a new gate, to be made of an also unspecified "low k" metal material. This Special K recipe entails a major manufacturing overhaul of the long dominant complimentary metal oxide silicon (CMOS) process, the ubiquity of which produced the volumes that made Moore's law and the semiconductor learning curve possible. Intel expects to be making chips with the new process by 2007, when geometries reach just 45 nanometers. But three years is a long time in silicon. Who knows what the dominant technologies in the industry will be at that time? Will the "value transistor" disrupt Intel's edge-of-the-envelope strategy? What about new power-management techniques and advances in traditional CMOS manufacturing?
See Intel's full presentation at:
http://www.intel.com/pressroom/archive/backgrnd/chau_high-k_metalgate_foils.pdf.

National Semiconductor (NSM)

SINGLE-CHIP SYSTEMS, ANALOG EXPERTISE,FOVEON IMAGERS

NOVEMBER 17: 40.14, 52-WEEK RANGE: 12.54 – 44.60, MARKET CAP: 7.44B

National now expects quarterly revenue to grow 7 – 10% sequentially, raising previous estimates of 4 – 7% growth over the fiscal Q1 total of $424.8 million. The company already had a large backlog of orders but said new "turns"—orders for delivery this quarter—have been stronger than expected. The next quarterly report comes December 4.

In product news, the company introduced the industry's fastest laser diode driver (LDD), used to read and write DVD and CD optical discs in PCs, laptops, and DVD players. The new LDD cuts the time it takes to "burn" a full 4.7-gigabyte DVD by 50%, from half an hour to 15 minutes. Sanyo was the first to incorporate the new chip into a DVD product. National also introduced five new low-power 10-bit and 12-bit analog-to-digital converters, key building blocks in a range of communications products like digital TV, high-definition TV receivers, base-station transceivers, communications receivers, data-acquisition systems, medical imaging, and consumer video.

Although NSM stock is up some 250% in the last year, it still trades at a price-to-sales ratio well below the industry average. Having shed two marginal divisions, moreover, its profit leverage going forward is better than ever.

Qualcomm (QCOM)

CDMA INTEGRATED CIRCUITS, IP, SOFTWARE

NOVEMBER 17: 46.72, 52-WEEK RANGE: 29.58 – 48.68, MARKET CAP: 37.38B

Third calendar quarter revenues were $909 million, and earnings were $291 million, up 53% year-over-year. The company ended the quarter with $5.4 billion in cash and short-term investments. It now sports a $38 billion market cap, with a seemingly rich P/E multiple of 47. But a constant flow of new innovations and products deepen the company's wireless dominance every day. The innovations, in turn, give legitimacy and leverage to its high-margin patent business. In the past month, Qualcomm:
• announced the RFR6500, a product that integrates multiple receive antennas on a single chip (two chips were previously required), boosting data performance and capacity of cdma2000 and 1x EV-DO networks;
• cooperated with China's Huawei in a successful test of 1x EV-DO in the 450 MHz frequency band;
• announced Sanyo and Toshiba selected the MSM6250 chipset for their WCDMA/GSM/GPRS multimode phones and that all 13 WCDMA infrastructure providers worldwide verified Qualcomm's MSM62xx line, the first products to achieve full interoperability across all WCDMA voice, circuit, and packet networks;
• said that most of the major mobile handset manufacturers—including Casio, Kyocera, LG Electronics, Samsung, Sanyo, Sony Ericsson, and Synertek—licensed Qualcomm's suite of mobile video capabilities: Qtv, Qcamcorder, and Qvideophone.

Synaptics (SYNA)

TOUCH-SENSORS, FOVEON IMAGERS

NOVEMBER 17: 12.73, 52-WEEK RANGE: 5.75 – 14.90, MARKET CAP: 305.95M

Third calendar quarter revenues were $29.6 million, up 33% from last year, and GAAP earnings were $2.3 million, or $.09 per share. The company reported a strong $19 million backlog for the fourth quarter and predicted a sequential revenue increase of 8-10%. It has $79 million in cash and short-term investments and no debt. With a price-to-sales ratio of 3, SYNA sells for less than the S&P 500 average.

Transmeta (TMTA)

MICROPROCESSOR INSTRUCTION SETS

NOVEMBER 17: 3.62, 52-WEEK RANGE: 0.91 – 5.51, MARKET CAP: 508.07M

Transmeta's financial performance still lags its innovative low-power technologies. September quarter revenues were just $2.7 million, down almost 50% sequentially and 58% year-over-year. Pro forma net loss was $18.7 million.

The company said its first-generation TM5800 chip, also known as "Crusoe," saw declining demand from notebook PC makers, as it gets long in the tooth. But Transmeta expects increasing use of the TM5800 from "ultra personal computer" makers and such devices as tablet PCs going forward.

The month's biggest news was the launch of the second-generation x86-compatible chip known as Efficeon. "Watt for watt," noted CEO Matthew Perry, "we believe our Efficeon processors perform much better than Intel's Pentium 4 and are competitive to Pentium M." He also said that initial customer feedback is positive. "We are seeing strong interest in Efficeon processors for designs beginning in the spring of 2004 in the mainstream consumer notebook segment. We are also receiving very strong feedback from leading notebook manufacturers for Efficeon processors based on 90 nanometer technology, which is expected to be in volume production in the second half of next year," Perry said.

In conjunction with the Efficeon unveiling, the company also gave an update on its LongRun2 technology, which is designed to mitigate the transistor leakage problems discussed above (see Intel update). Calling LongRun2 "a major breakthrough in energy conservation," Perry told investors the company had "demonstrated an experimental version of our Efficeon processor that adjusted leakage up to hundreds of times per second while playing a video game, playing a DVD movie and going into standby mode…. In standby mode, Efficeon's core leakage power was reduced by approximately 70 times when using the LongRun2 technology."

The company expects December quarter revenues to increase anywhere from zero to 50%, depending on how fast Efficeon sells. It projects cash holdings of $52 million at year end.

Xilinx (XLNX)

PROGRAMMABLE LOGIC DEVICES

NOVEMBER 17: 35.06, 52-WEEK RANGE: 18.50 – 35.20, MARKET CAP: 12.00B

The company announced it shipped more than 6 million of its low-cost, high-volume Spartan FPGAs in the September quarter. Spartan was the first FPGA to be produced using 300 mm wafers and 90 nm process, which yields more than 5 times the number of chips per wafer compared to the 200 mm wafer/130 nm process combination. The resulting lower costs allow FPGAs to move into markets previously served by ASICs. More than 70 million Spartan chips have been sold since their introduction in 1998.

Like Altera, Xilinx is pushing its serial connectivity solutions. The company shipped more than 200,000 RocketIO multi-gigabit transceivers in the September quarter, more than 10 times that of its "nearest competitor." With Xilinx marketing its own lightweight, link-layer serial protocol named Aurora, and Altera announcing SerialLite, the two companies are engaged in an intense competition for high-speed serial applications.

CONTINUED FROM PAGE 3

synonymous.

As the design emphasis shifts from cost performance to cost-performance-per-watt, the microprocessor will lose its 30-year preeminence as the workhorse of embedded systems.

In another coming-in-under-the-radar story, the programmable logic device began as an array of transistors that would be interconnected, after manufacture, in a specific pattern to consolidate "glue logic" in systems. Glue logic is the miscellaneous logic that ties big chips together. But although PLDs waste transistors, as PLD capability grew, PLDs overtook low-end applications that belonged to ASICs and ASSPs. Now PLDs are appearing in applications that traditionally belonged to microprocessors and DSPs. ASICs and ASSPs will continue to decline as PLDs take over their applications.

Foundries, such as Chartered, Taiwan Semiconductor, and United Micro, will increasingly dominate chip production

The future belongs to untethered systems. Designers implement today's untethered systems with some combination of ASICs, ASSPs, microprocessors, and DSPs. The future's untethered systems will demand the flexibility to adapt to evolving standards and to available wireless protocols. The new wireless standards also demand more performance. The rising cost of ASIC development, combined with the ASIC's inflexibility, makes ASICs and ASSPs less suitable for untethered applications. Microprocessors and DSPs, which dominated tethered applications, can't meet the cost-performance-per-watt requirements of untethered systems. That leaves PLDs. But today's PLDs from **Altera** (ALTR) and **Xilinx** (XLNX) are also unsuitable for untethered applications. That leaves no candidates to exploit growing opportunities in untethered systems.

Why it's happening

ASIC design starts are decreasing because chip-design costs are too high, because mask costs are too high, and because programmable logic devices are adequate for more and more applications. ASSPs will decline for similar reasons. The mask set for a chip in a 130-nm process can cost $650,000; the mask set for a chip in a 90-nm process can cost $1,200,000. And if the first chip isn't perfect, you'll be writing an identical check for the second mask set. Mask costs double with each process generation. Semiconductor-process development is also expensive, at $500 million or so for each process generation. In addition, chip designers need new, more expensive software to design the chips of each new process generation. As a rule of thumb, chip design costs are about ten times mask costs. If total chip costs double, the market has to double to keep the product's price from rising. Each process generation thus narrows the range of profitable applications. This means that chips made in advanced processes must be either high volume or high margin.

Beyond high costs, I think there's a more fundamental reason for the decline in ASICs and ASSPs. These components aren't a good fit for today's design methods. To make money, you have to get the custom chip right the first time and that's just about impossible when hundreds of millions of logic transistors are involved. Make the smallest error and there are huge penalties in dollars, in engineering time, and in delayed market entry.

Leading-edge ASICs can achieve about ten times the performance and about twenty times the circuit capacity of leading-edge PLDs. "Structured ASICs," from companies such as **Chip Express**, **eASIC**, and **Lightspeed**, attempt to fit into the performance and capacity gap between ASICs and PLDs by making chips that are mostly standard, but with a few chip layers reserved for customization. This makes the chips more general and spreads the mask cost across several applications. But from the point of view of the designer, the structured ASIC's design looks more like ASIC design than it looks like PLD- or microprocessor-based design, relegating it to a segment of declining interest. Lightspeed recently announced that it will leave the chip business in favor of a licensing model.

ASIC design methods try to achieve a perfect design the first time, every time, but engineers don't work that way. In microprocessor-based systems, for example, they build general-purpose hardware and write software for each application in what becomes a series of program-run-debug experiments. They get something working and correct errors until the system does what they want. That's the opposite of ASIC design, but it's a match for PLDs. That's the primary reason that PLDs are displacing ASICs and ASSPs.

It is also a match for microprocessors and DSPs, which are thoroughly entrenched using the mainstream engineering design method: problem solving by writing software programs. That fact makes encroachment by PLDs seem unlikely because PLD-based implementation requires hardware-design skills. But in untethered systems microprocessors and DSPs are too energy inefficient to dominate.

When the race began, ASICs took the lead in high-volume, performance-oriented systems, from graphics accelerators to routers. A custom circuit has the best performance and, for high-volume applications, it has the lowest cost per chip because the custom circuit is the smallest, most direct implementation of a particular function. But as ASIC costs rose and their markets narrowed, ASSPs grew into applications where similarly escalating ASSP costs could be amortized across several end-product suppliers. PLDs took the lead in consolidating glue logic and slowly gained ground. Microprocessors took the lead in cost-oriented systems where their performance was adequate. DSPs gained in applications that needed more number-crunching performance than a general-purpose microprocessor could provide. Microprocessor and DSP applications grew more rapidly than ASIC, ASSP, and PLD applications because their software programming model enabled a larger pool of designers. Thus, cost and performance sorted the contestants into particular applications.

Now the rules are changing. Power and flexibility are becoming more important. ASICs and ASSPs lack flexibility. Microprocessors and DSPs lack energy efficiency. Today's PLDs lack both performance and energy efficiency. That leaves no candidate to win the logic-mile event.

Consequences

Leading ASIC suppliers, such as **Agere Systems** (AGRb), **Agilent** (A), **LSI Logic** (LSI), and **Toshiba** (TOSBF.PK), will struggle as ASIC design starts continue to decline. ASSP suppliers, such as **Analog Devices** (ADI), **Qualcomm** (QCOM), and Texas Instruments, will have success in untethered systems until better alternatives appear. Microprocessors will move from workhorse to supervisor in untethered systems. Altera and Xilinx will grow into applications that once belonged to ASICs and ASSPs, but will have some difficulty invading microprocessor and DSP applications. As PLDs improve and as their development software improves, PLDs will begin to displace DSPs. PLDs are already beginning to replace most of the DSPs in what were DSP-based computational arrays.

PLD suppliers Altera and Xilinx look like fabless chip companies, but they are really software companies. Their hold on the market comes from their design tool software. The largest chip makers, **Advanced Micro Devices** (AMD), **AT&T** (T), **IBM** (IBM), Intel, and **Motorola** (MOT), took them on and failed. The largest design tool software companies such as **Cadence** (CDN) and **Synopsys** (SNPS) thought the PLD players were chiefly chip companies and never took them on. Chip sales pay for continued chip development, but chip sales also subsidize software development and the expansion of intellectual property (IP) libraries. Unlike microprocessor suppliers who are forced to maintain instruction-set compatibility from generation to generation, PLD makers sell a component that is circuit-generic in manufacture and circuit-customized in the field. PLDs will adopt new logic elements to suit emerging applications and they will adopt new non-volatile memory cells for personalization memory.

The memory hurdles

What's happening

Flash memory, DRAM (dynamic random-access memory), and SRAM (static random-access memory) are the long-time leaders in memory chips. Each occupies a niche in the personal computer. Their markets grew with the PC market. But the design emphasis is changing from the mostly tethered PC to untethered systems.

What components will fill the memory sockets in untethered systems? Flash memory, DRAM, and SRAM, alone or in combination, have serious shortcomings. Untethered systems need non-volatile memory that's as dense as DRAM and is as fast as SRAM. Flash memory is too slow and it wears out. DRAM and SRAM can't retain information through power cycles and SRAM burns too much power. Combinations of memory types take too many sockets, cost more than a single component, and burn too much power. There's opportunity for a new memory type to replace all three incumbents.

Why it's happening

The memory event began with read-only memory (ROM) and DRAM. Flash memory soon displaced ROM because, with flash memory, stored values could be occasionally updated in the field. When the personal computer was introduced in 1981, the time to access memory matched the instruction time of the microprocessor. Since that time, however, memory speeds and microprocessor speeds diverged because memory designers optimized for capacity and microprocessor designers optimized for speed. The result is that today's leading-edge DRAMs have 4,000 times their 1981 capacity, but are only five to seven times faster. Meanwhile, microprocessors operate at 600 times the PC's original rate.

With microprocessors issuing several instructions per clock tick, the gap between the performance of microprocessors and DRAM is even wider than the ratio of their clock speeds. SRAM memory cells built with the speed of microprocessor circuits entered the race to bridge the speed gap between DRAMs and microprocessors.

In twenty years of development, flash memory, DRAM, and SRAM components locked up niches in the PC. Made chiefly by Intel, AMD, Toshiba, Sony, and **Atmel** (ATML), flash memory, which retains information as power is cycled, holds the code that initializes the chips on the main system board. Made chiefly by **Samsung** (05930.KS) and Micron, DRAM, which has the highest capacity, is the system's working memory. Made chiefly by **Cypress** (CY), Micron, Samsung, and the new **Renesas** (Hitachi and Mitsubishi Electric partnership), SRAM is fast and expensive. Each has advantages that assure its PC niche. Each has flaws that prevent it from invading niches held by the other memory types. Flash memory is slow and it wears out. DRAM loses its contents when power is off. SRAM lacks the density of DRAM and it uses lots of power.

Back when no PC's performance was good enough, consumers bought high-margin, leading-edge PCs as soon as they came on the market. But when many PCs' performance exceeds users' demand, consumers buy "value PCs" and purchasing cycles lengthen. Margins in the PC business decline. With the emergence of the low-margin value PC, the industry's design resources are being reallocated to higher-margin untethered systems. The move to untethered systems changes the design goal from cost performance to cost-performance-per-watt.

Consequences

Flash memory, DRAM, and SRAM will hold their positions in the PC and will continue to gain in applications where they have an advantage and where their flaws aren't crippling. One example is smart cards, where flash memory is needed as non-volatile storage and where its slow speed and limited life aren't serious problems. Continuing opportunities in flash memory, DRAM, and SRAM include consumer appliances in emerging economies such as China. Other leading makers and module designers for flash memory, DRAM, and SRAM include **Hynix** (HXSEY.PK), **Infineon** (IFX), and **SanDisk** (SNDK).

Non-volatile memories have been around for fifteen years and high-volume PC production ensured their lower costs. No new memory candidate could compete on cost, so none has entered high-volume production. The leading candidates in non-volatile memory are ferroelectric random-access memory (FRAM), magnetoresistive random-access memory (MRAM), and ovonic unified memory (OUM). FRAM's backers include Agilent, **Hitachi** (HIT), IBM, Infineon, Micron, Motorola, **NEC** (NIPNY), **Ramtron** (RMTR), Samsung, and Texas Instruments. MRAM's backers include Cypress Semiconductor, Hitachi, IBM, Infineon Technologies, **Mitsubishi**, Motorola, NEC, **Philips** (PHG), Samsung, **STMicroelectronics** (STM), Toshiba, Taiwan Semiconductor, and **Union Semiconductor**. OUM's backers include **Azalea Microelectronics**, **BAE Systems PLC** (BAESF.PK), Intel, **Ovonyx** (Energy Conversion Devices spin-off), **STMicroelectronics**, and Toshiba. And startup **Axon Technologies** is developing a brand new type of non-volatile memory called Programmable Metalization Cell memory (PMCm). There is no clear leader and there are many contenders.

Shifting design resources from PCs to untethered systems increases the incentive to develop fast, dense, non-volatile memories. While the PC dominated semiconductor memory sales, each incumbent memory type held its cost-sensitive niche through volume PC production. Today's memory chips will hold temporary sockets in untethered systems, but they cannot lock them in because they all have major disadvantages for untethered applications. The sockets will wait for the winner of the non-volatile memory competition.

Wrap-up

The value PC and the value transistor are transforming the semiconductor industry. It's difficult to see the transformation because the changes aren't abrupt. It's not as if one day there's no value transistor and the next day there is one and everything is different. The value transistor is like a tide rising on the applications beach. Some applications are above the tide and some applications are below the tide. There's a similar story for the value PC.

The value transistor levels the playing field for semiconductor competitors, because how transistors are used is becoming more important than how small they are. The value transistor means more innovation because the barrier to participation is lower (no fab needed). The value transistor means most fabs are good enough. Fabs that are good enough decrease the advantage of being an integrated device manufacturer and they decrease the need for leading-edge semiconductor-processing equipment.

As a result of the shift to untethered systems a new memory type will emerge. The PC's entrenched memories (flash memory, DRAM, and SRAM) are all unsuitable for untethered systems. The new memory will have the non-volatility of flash memory, the density of DRAM, and the speed of SRAM.

The new memory that emerges for untethered systems will greatly benefit PLDs, which today use bulky, expensive, power-hungry SRAM for personalization memory. The new memory, combined with improved logic elements, improved interconnect, and improved development tools, will make PLDs fit untethered applications.

-Nick Tredennick and Brion Shimamoto

GILDER TECHNOLOGY REPORT PUBLISHED BY GILDER PUBLISHING, LLC AND FORBES INC.

291A MAIN STREET, GREAT BARRINGTON, MA 01230, TEL: (413) 644-2100, FAX: (413) 644-2123 • EMAIL: INFO@GILDER.COM