a stable product architecture. In this context, learning means learning about components and the core concepts that underlie them. Given the way knowledge tends to be organized within the firm, learning about changes in the architecture of the product is unlikely to occur naturally. Learning about changes in architecture—about new interactions across components (and often across functional boundaries)—may therefore require explicit management and attention. But it may also be that learning about new architectures requires a different kind of organization and people with different skills. An organization that is structured to learn quickly and effectively about new component technology may be ineffective in learning about changes in product architecture. What drives effective learning about new architectures and how learning about components may be related to it are issues worth much further research.

These ideas also provide an intriguing perspective from which to understand the current fashion for cross-functional teams and more open organizational environments. These mechanisms may be responses to a perception of the danger of allowing architectural knowledge to become embedded within tacit or informal linkages.

To the degree that other tasks performed by organizations can also be described as a series of interlinked components within a relatively stable framework, the idea of architectural innovation yields insights into problems that reach beyond product development and design. To the degree that manufacturing, marketing, and finance rely on communication channels, information filters, and problem-solving strategies to integrate their work together, architectural innovation at the firm level may also be a significant issue.

Finally, an understanding of architectural innovation would be useful to discussions of the effect of technology on competitive strategy. Since architectural innovation has the potential to offer firms the opportunity to gain significant advantage over well-entrenched, dominant firms, we might expect less entrenched competitor firms to search actively for opportunities to introduce changes in product architecture in an industry. The evidence developed here and in other studies suggests that architectural innovation is quite prevalent. As an interpretive lens, architectural innovation may therefore prove quite useful in understanding technically based rivalry in a variety of industries.

CASE II-12

Intel Corporation: The DRAM Decision

George W. Cogan and Robert A. Burgelman

INTRODUCTION

In November 1984, *Andy Grove*,[1] Intel's chief operating officer, stood in his office cubicle gazing out at Silicon Valley and thought about his company's future. The semiconductor industry which Intel had helped create 16 years earlier had entered what looked to be a prolonged cyclical downturn. Some operations had already been trimmed, but Grove believed the company would have to react again soon (see company financial data in Exhibit 2). The recession hit the company's Memory Components Division particularly hard. For much of the previous five years, memory components had been suffering under competitive pressure from the Japanese.

Since 1980, Intel had been losing its market position in *dynamic random-access memories (DRAMs)* as the industry average selling price per chip had declined much more rapidly than the 20 to 30 percent per year which was customary. The Japanese had taken the lead in unit sales of the latest generation of DRAMs, the 256 *kilobit* (256K) version, but Intel was fighting back with a program to leapfrog the Japanese in the product's next generation. Its $50 million 1 *megabit* (1 meg = 4 × 256K) research project was soon to produce working prototypes. Intel managers estimated they were ahead of the Japanese in the 1 meg device. Still, a debate was growing within the company about whether Intel could continue to compete in the commodity market of DRAMs. Grove was formulating his personal position on the matter.

It seemed clear that if Intel chose to continue with the DRAM product line, it would have to commit to at least one $150 million state-of-the-art *Class 10 production facility*. On the other hand, Intel's other businesses were much more profitable than memories; in an ROI framework, the microprocessor business deserved the majority of Intel's corporate resources. It was difficult for both Grove and *Gordon Moore*, Intel's chief executive officer, to imagine an Intel without DRAMs. The memory business had made Intel, and was still by far the largest market segment in integrated circuits. Not the least of Grove's worries was how the investment community would react to Intel's decision to cede such a large market segment to the Japanese.

[1] Note: All italicized names appear with biographies in Exhibit 1; all italicized words appear with definitions in the technical appendix.

EXHIBIT 1 Biographies of Key Intel Personnel

Jack Carsten joined Intel from Texas Instruments and has held various high level management positions since then. In 1985 he was senior vice president and general manager of the Components Group.
Dennis Carter is a Harvard M.B.A. with an engineering background. He has worked in several areas of the company and is currently assistant to the president.
Sun Lin Chou received his B.S. and M.S. degrees in Electrical Engineering from MIT and his Ph.D. in electrical engineering from Stanford University. He joined Intel in 1971 and has managed the DRAM technology development group in Oregon since then.
Dov Frohman joined Intel from Fairchild in 1969. He was responsible for the invention of the EPROM. He currently manages Intel's design group in Israel.
Edward Gelbach joined Intel from Texas Instruments in 1969. He is currently senior vice president of sales.
Andrew Grove was born in Budapest. He received his B.S. from CCNY and his Ph.D. from Berkeley. After working at Fairchild Camera and Instrument for five years, he joined Intel in 1968. He has been president and chief operating officer since 1979.
Ted Hoff joined Intel as a designer in 1969. He headed the group that invented the microprocessor. Hoff left Intel in 1983.
Gordon Moore was born in San Francisco. He received his B.S. in chemistry from Berkeley and his Ph.D. in chemistry and physics from the California Institute of Technology. He worked as a member of the technical staff at Shockley Semiconductor from 1956 to 1987, and he founded Fairchild. He founded Intel in 1968 and is currently the chairman and CEO.
Robert Noyce was born in Burlington, Iowa. He received his B.S. from Grinnell College and his Ph.D. from MIT. He was a research engineer at Philco from 1953 to 1956, a research engineer at Shockley Transistor, and a founder and director of Fairchild Camera and Instrument. He is credited with coinventing (with Kilby at TI) the integrated circuit. He founded Intel and currently serves as vice chairman of the board of directors.
Bob Reed received his bachelor's degree from Middlebury College and his M.B.A. from the University of Chicago. He joined Intel in 1974. He was appointed chief financial officer in 1984.
Ron Smith received his bachelor's degree in physics from Gettysburg College and his M.S. and Ph.D. degrees in physics from the University of Minnesota. He joined Intel in 1978 as a device physicist in the Static Logic Technology Development Group. In 1985, he was manager of that group.
Dean Toombs joined Intel from Texas Instruments in 1983 with the express purpose of running the Memory Components Division.
Leslie Vadasz joined Intel in 1968 and has held a variety of senior management positions since then. He is currently senior vice president and director of the Corporate Strategic Staff.
Ron Whittier holds a Ph.D. in chemical engineering from Stanford University. He joined Intel in 1970. From 1975 until 1983, he managed the memory products division. In 1983, he became vice president and director of Business Development and Marketing Communications.
Albert Yu was born in Shanghai and holds a Ph.D. in electrical engineering from Stanford University. He joined Intel in 1975.

COMPANY BACKGROUND

Early History

On August 2, 1968, the *Palo Alto Times* announced that *Bob Noyce* and *Gordon Moore* had left Fairchild to form a new company. Andy Grove, who had been Moore's assistant director of research at Fairchild, also left to complete what the company's historians have called the triumvirate. The three were key technologists in the emerging solid-state electronics industry. Noyce had invented the integrated circuit (simultaneously with Jack Kilby at Texas Instruments), and Intel was the first company to specialize in making large-scale integrated circuits.

In mid-1969, Intel introduced its first product, a *bipolar static random-access memory (SRAM)* with a 64-bit storage capacity. The chip itself was less than a quarter of an inch on a side and contained nearly 400 *transistors*. While the SRAM had some small markets, Intel had set its sights on the growing computer memory business, then dominated by *magnetic core* technology. To attack the magnetic core business required at least a 10-fold reduction in cost per bit.

The Intel managers decided early on to pursue a new process technology in addition to the relatively proven bipolar process. The *metal-oxide-semiconductor* (MOS) process promised to lead to increased transistor density while simultaneously reducing the number of fabrication steps required to make a working chip. The process had been published in scientific journals, but serious manufacturability questions remained. MOS transistors consumed only a fraction of the power of a traditional bipolar transistor and thus could be more densely packed on the chip. But they were also very sensitive to trace amounts of impurities in processing, raising the question

EXHIBIT 2 Selected Intel Corporation Financial Data

	Year ended December 31								
	1976	1977	1978	1979	1980	1981	1982	1983	1984
Sales	226	283	400	663	854	788	900	1,122	1,629
COGS	117	144	196	313	399	458	542	624	883
Gross margin	109	139	204	350	455	330	358	498	746
R&D	21	28	41	67	96	116	131	142	180
SG&A	37	48	76	131	175	184	198	217	315
Operating profit	51	63	87	152	184	30	29	139	251
Interest and other			(1)	(3)	2	10	2	40	47
Profit before tax	51	63	86	149	186	40	31	179	298
Income tax	26	31	42	71	89	13		63	100
Net income	25	32	44	78	97	27	31	116	198
Depreciation	10	16	24	40	49	66	83	103	114
Capital invest	32	97	104	97	152	157	138	145	388

	December 31								
	1976	1977	1978	1979	1980	1981	1982	1983	1984
Cash and ST invest	26	39	28	34	127	115	85	389	230
Working capital	93	81	67	115	299	287	306	608	568
Fixed assets	30	80	160	217	321	412	462	504	778
Total assets	156	221	356	500	767	871	1,056	1,680	2,029
LT debt	0	0	0	0	150	150	197	127	146
Equity	109	149	205	303	432	488	552	1,122*	1,360
Employees	7,300	8,100	10,900	14,300	15,900	16,800	19,400	21,500	25,400
ROS	11.1%	11.3%	11.0%	11.8%	11.4%	3.4%	3.4%	10.3%	12.2%
ROA**	24.3%	20.5%	19.9%	21.9%	19.4%	3.5%	3.6%	11.0%	11.8%
ROE**	33.8%	29.4%	29.5%	38.0%	32.0%	6.3%	6.4%	21.0%	17.6%

Note: The first and second quarters of 1985 showed revenue of $375 million and $360 million and profit of $9 million and $11 million, respectively. The first and second quarters of 1984 showed revenue of $372 million and $410 million and profit of $54 million and $50 million, respectively.
*Includes $250 million proceeds from sale of 11% stake to IBM.
**Based on beginning-of-year asset (equity) values
Source: Intel annual reports.

of whether their performance characteristics would remain stable over time.

Les Vadasz headed the MOS team of several engineers. In contrast to the bipolar effort, the MOS effort moved slowly. The primary problem was to develop a stable transistor *threshold voltage*. After a year of frustration and setbacks, Vadasz's team produced the first commercially available MOS SRAM, the 256-bit "1101." The successful processing sequence had several proprietary aspects which put Intel in the forefront of semiconductor technology development. Vadasz commented that at this early stage of development, the processing sequences had proprietary aspects, but were not always well understood.

Since the market for SRAMs was young, Intel had difficulty selling the new device. But the successful MOS process was immediately applied to the existing market for *shift registers* among mainframe computer makers. Shift register sales provided the company with a war chest of cash needed to weather its first semiconductor recession of 1970–1971.

Development of DRAM

Another technical innovation followed the 1101. Intel worked closely with Honeywell engineers to design and develop the first DRAM in 1970, the 1-kilobit "1103." While the SRAM required six MOS transistors per memory cell, the DRAM required only three transistors. With fewer elements in each memory cell, the 1103 contained more storage capacity in the same silicon area. While the new design allowed increased memory cell density, it also required a significant amount of external circuitry for *access* and *refresh*. An advertisement placed in computer trade journals in early 1971 announced: "THE END. CORES LOSE PRICE WAR TO NEW CHIP."

In spite of the price/performance advantage, customers had to be taught how to use the new device and

convinced of its reliability. *Ed Gelbach*, VP of sales, remembered 1971:

> We could never find a customer that used them and yet we were shipping literally hundreds of thousands of them. They were all testing them and putting them in boards . . . but it seemed like none of the customers ever shipped machines with the part. My recurring nightmare was that all of those chips would be returned over a single weekend.

In order to speed the adoption of DRAMs, Intel started the Memory Systems Operations (MSO), which assembled 1103 chips along with the required peripheral controller circuitry for OEM sale into the computer maker market. Soon MSO was responsible for about 30 percent of Intel's business. By 1972, the 1103 was the largest selling integrated circuit in the world and accounted for over 90 percent of Intel's $23.4 million in revenue.

Gordon Moore called the 1103 "the most-difficult-to-use semiconductor product ever invented." Ironically, that may have helped its market success:

> There was a lot of resistance to semiconductor technology on the part of the core memory engineers. Core was a very difficult technology and required a great deal of engineering support. The engineers didn't embrace the 1103 until they realized that it too was a difficult technology and wouldn't make their skills irrelevant.

New DRAM Generations

From its early days, Intel was fighting a battle with processing yields. The early 1103's were produced on 2-inch-diameter silicon wafers, each containing about 250 devices. Of the 250, early 1103 runs produced an average of 25 fully functional devices, or an overall yield of 10 percent. *Ron Whittier*, general manager of the Memories Components Division from 1975 until 1983, said that throughout a product's life cycle, wafer yields increased continually as process improvements were developed. The productivity of the factory was also increased by changing the size of the wafer whenever silicon manufacturers developed techniques to grow larger silicon ingots and equipment manufacturers developed machines which could handle larger wafers. In 1972, *Albert Yu* headed a team which converted the bipolar process from 2-inch to 3-inch wafers, effectively doubling capacity.

In the early days, Vadasz recalled, MSO developed another strategy for increasing yield. Since it only took one defective memory cell (out of 1024 in the 1103) to make a chip dysfunctional, it seemed inefficient to throw away all defectives. MSO's scheme was to compensate for a defective memory cell using creative peripheral logic circuitry. The peripheral circuitry was designed to bypass the defective cells within each memory chip so that rejected 1103s could still be used. Since the scheme required extra 1103s in each system, Intel referred to the concept as redundancy.

Soon after Intel's early success, competitors entered the market for DRAMs and began to erode Intel's MOS process technology lead. By the mid-1970s, Intel was one of several companies vying to be the first at introducing the new generation of DRAM memories. Every three years, a new generation with four times as much capacity as its predecessor was developed (see Exhibit 3).

Vadasz recalled that even at the 4K and 16K level, Intel was struggling to keep up with its competitors. During the formative years of the DRAM market, the chip design was in rapid flux. A start-up company, MOSTEK, was able to take market share from Intel in the 4K generation by incorporating the peripheral circuitry required to manage the memory on the chip itself. Vadasz recalled: "The first DRAMs were not very user-friendly, and MOSTEK came out with a better product." MOSTEK introduced the concept of on-chip *multiplexing*, which allowed a smaller number of output pins to address the entire memory. Multiplexing started a trend in DRAMs towards user-friendliness.[2]

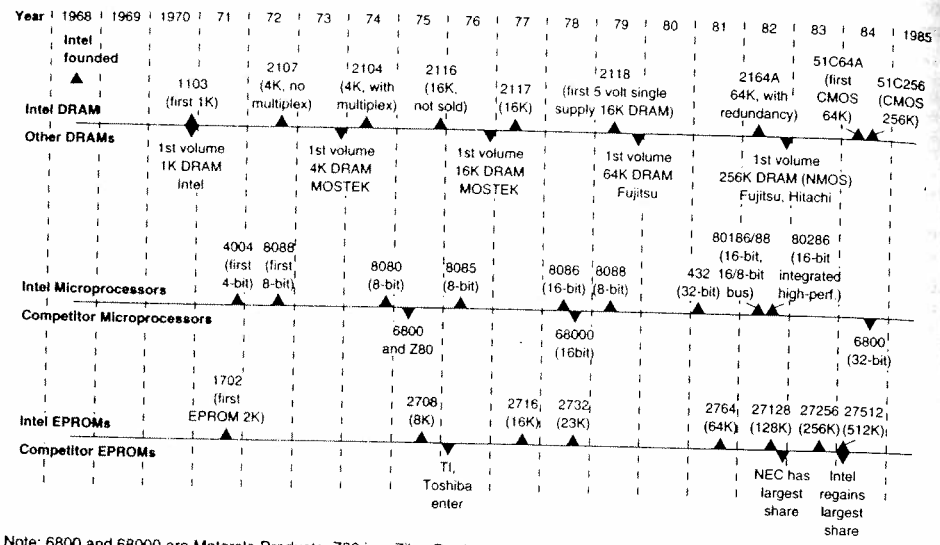
Vadasz commented:

> Even though you have invented the product, sometimes it is easier for new entrants to seize an opportunity and beat you to the punch. They are not encumbered by the same things you are. . . . The real problem in technological innovation is in anticipating the relevant issues. Once a technological "box" has been defined, it is easy for a team of great engineers to optimize everything in that box. Choosing the box is the hard part.

Intel's first 4K DRAM was redesigned to include the internal multiplexing logic. *Sun Lin Chou*, who was involved in the 4K DRAM development, said that in the revised version, Intel also implemented a one-transistor DRAM cell, which became the industry standard. While more challenging from a process technology standpoint, the reduction in the number of transistors allowed for a smaller chip size. The revised 4K version sold well, but time was short before the next generation.

[2]Eventually Intel sold MSO since the value added had been integrated onto the chip itself and the majority of MSO's customers had learned how to use DRAMs.

EXHIBIT 3 Product Introduction Timelines



Note: 6800 and 68000 are Motorola Products. Z80 is a Zilog Product.
Source: Intel documents, Dataquest.

Dennis Carter described Intel's early strategy as "staying ahead of the experience curve using process technology."[3] According to Sun Lin Chou, a successful DRAM company participates in the early phase of each generation when low competitor yields and high demand support high prices.

> In fact, for the first two years, the demand for DRAMs to the first market entrant is semi-infinite. As soon as the leading vendor makes a new DRAM, he can crank his capacity to the maximum and he will be guaranteed of selling all his output. This is not true for more complex products such as a logic product with a new function where the customers have to first learn how to use it.

Each new generation required a quadrupling of the number of transistors contained on a chip. The driving force behind increased density was the ability to define patterns of ever narrower dimensions (functional equivalent of wires and components in a circuit) on the silicon wafer, to invent creative ways of reducing the required number and size of components per memory cell, and to make larger chips without defects. Each new generation reduced the minimum linewidth by a factor of about 0.7, from 5 μm at the 4K generation. The minimum linewidth was controlled primarily by the accuracy of the photolithography process, while the maximum chip size was determined by the ability to control the number of random defects on the wafer.[4]

While competition was tough even at the 4K level, a series of process innovations kept Intel amongst the memory leaders through the 16K DRAM generation (see Exhibit 3). Gordon Moore developed the strategy of using DRAMs as a technology driver. The latest process technology was developed using DRAMs and later transferred to other products. Early on in the company's development, Intel managers decided to merge the research and manufacturing functions. Gordon Moore had been dissatisfied with the linkage between research and manufacturing at Fairchild. As a result, he had insisted that Intel perform all process research directly on the production line. Moore commented:

> Our strategy optimizes our ability to make fast incremental process technology improvements. We don't have a central corporate research lab. We tend to evaluate other research advances in light of how they will affect our businesses. For instance, while Texas Instruments has been funding a research effort in *gallium arsenide*, we have been watching gallium arsenide develop for the past 20 years. We're still silicon believers.

During the 1970s, Intel competed by developing new processes which were used to enhance product features or to enable new product families beyond memories. The *HMOS* (high-performance MOS) process enabled Intel to introduce the first 5-volt-single-power-supply 16K DRAM in 1979. Earlier offerings, including Intel's two previous 16K DRAMs (2116 and 2117), required that the user supply three separate voltages to the chip. The new product, the 2118, greatly simplified the user's design and production tasks. While Intel had lost market share with the 2116 and 2117, it was all alone with the 5-volt device and captured a price premium of double the industry average for three-power-supply 16K DRAMs in 1979 (see Exhibit 4). The DRAM technology development group focused a significant amount of its resources on developing Intel's third 16K DRAM offering while competitors concentrated on the 64K generation.

Intel management decided to focus on the single-power-supply 16K DRAM for two primary reasons: they projected a relatively long life cycle for the 16K generation due to the technical challenge in achieving the 64K generation, and they believed the one-power-supply process would eventually dominate the memory industry. They considered it too risky to tackle both the 64K DRAM generation and the single-power-supply technology in the same product.

The drive towards smaller and smaller geometries was achieved through improvements in both processing methodology and processing machinery. Dennis Carter explained that in the early years some processing steps were considered black magic and defined a company's competitive edge. As time went on, the movement of engineers between chip companies and the involvement of suppliers and equipment manufacturers in process development efforts led to a general leveling of process capability amongst Silicon Valley firms. Sun Lin Chou commented about the trends in processing:

> Process technology and equipment have become so complex and expensive to develop that no vendor can hope to do better than [its] competitors in every process step. The key to innovation is to be on par with your competitors on every process step, but to select one or two or three process features with the highest leverage and focus your efforts to gain leadership there. In DRAMs we focus on high-quality thin *dielectrics*.

The Invention of the EPROM

Albert Yu, vice president of development and general manager of the components division, said he usually associates the invention of any important product with one person. The EPROM (electrically programmable read-only memory) was invented by *Dov Frohman*. Yu said Frohman not only invented the product, but he also described the physical effect, saw that it could be applied to a memory device, designed the first part, and fabricated the first device.

Frohman's story has become legendary at Intel. As a recent hire from Fairchild in 1969, Frohman was assigned to help understand and remedy a strange phenomenon which was causing reliability problems with the MOS process. The problem involved the silicon gate structure. Frohman saw that the phenomenon could be explained by the existence of an unintentional *floating gate* within the MOS device. He realized that if a floating gate were intentionally constructed, a new type of programmable memory which would permanently store information could be built.

Frohman designed the first test devices and assembled a demonstration for Gordon Moore. According to Frohman:

> We put together a 16-bit array with primitive transistor packages sticking out of the 16 sockets, an oscilloscope and pulse generator, and we carted all this into Gordon's office. There were red bulbs to indicate the bits. This was all new to us, and we were thrashing around. We showed Gordon that by pushing the button you could program the device, and we demonstrated that it would hold a charge.

Later, it was discovered that ultraviolet light could be used to erase the memory. Moore committed the company to the production of the EPROM even though no one could tell where the device would have applications. Recalled Moore:

[3]The experience curve referred to the declining nature of industry-wide manufacturing costs over time due to experience. The semiconductor industry had a 70 percent experience curve (costs reduced by 30 percent for each doubling in cumulative volume). Companies who were not ahead of the curve for a particular product or generation suffered erosion of margins or market share.

[4]The size of the chip defines the area of the wafer which each process-induced defect can potentially damage. If the chip size is too large, yields are unacceptably low unless the defect level can be simultaneously reduced.

EXHIBIT 4 Market Information for DRAMs and Microprocessors

Product	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
Worldwide unit shipments of DRAMS (in thousands)											
4K	615	5,290	28,010	57,415	77,190	70,010	31,165	13,040	4,635	2,400	2,250
16K 3PS*			50	2,008	20,785	69,868	182,955	215,760	263,050	239,210	120,690
16K 5V*						150	1,115	5,713	23,240	57,400	40,600
64K					1	36	441	12,631	103,965	371,340	851,600
256K									10	1,700	37,980
Worldwide yearly average selling prices of DRAMs ($/unit)											
4K	17.00	6.24	4.35	2.65	1.82	1.92	1.94	1.26	1.62	2.72	3.00
16K 3PS*			46.39	18.63	8.53	6.03	4.77	2.06	1.24	1.05	1.09
16K 5V*						17.67	7.38	3.84	2.23	1.98	2.07
64K					150.00	110.14	46.26	11.00	5.42	3.86	3.16
256K									150.00	47.66	17.90
Total market	10,455	33,010	124,163	189,559	317,932	562,339	961,785	621,775	905,506	1,885,745	3,593,242
Intel DRAM market share											
4K	82.9%	45.6%	18.7%	18.1%	14.3%	8.7%	3.2%				
16K 3PS*			37.0%	27.9%	11.5%	4.4%	2.1%	2.4%	2.3%	1.9%	1.4%
16K 5V*						100.0%	94.0%	66.5%	33.1%	11.7%	12.3%
64K							0.7%	0.2%	1.5%	3.5%	1.7%
256K											0.1%
Estimated revenue**	8,667	15,052	23,643	37,976	40,479	32,882	28,139	25,534	33,109	68,238	58,607
Microprocessor sales history by architecture											
Achitecture (% units sold)			**1976**	**1977**	**1978**	**1979**	**1980**	**1981**	**1982**	**1983**	**1984**
8-bit:											
Zilog (780)			2.2%	5.8%	12.4%	17.0%	21.1%	22.7%	23.4%	37.4%	35.1%
Intel (8080, 8088)			22.8%	36.6%	34.6%	38.9%	27.1%	19.7%	19.1%	22.3%	33.5%
Motorola (6800, 650X, 680X)			15.0%	13.0%	17.2%	20.8%	18.8%	21.1%	17.9%	14.8%	14.0%
Others			60.0%	44.6%	35.8%	23.3%	33.0%	36.5%	39.6%	25.5%	17.4%
Total 8-bit (million units)			n/a	n/a	n/a	12.5	22.4	33.8	47.9	67.8	75.1
Average selling price			n/a	n/a	n/a	$6.03	$4.60	$3.32	$3.18	$3.25	$4.06
16-bit:											
Zilog (Z8000)						1.4%	4.5%	3.4%	5.1%	5.8%	6.2%
Intel (80186/286,8086)					6.7%	14.0%	28.9%	31.7%	26.6%	32.1%	59.1%
Motorola (68000)								3.9%	5.8%	10.8%	20.2%
Others					93.3%	84.6%	66.6%	61.0%	62.5%	51.3%	14.5%
Total 16-bit (million units)			n/a	n/a	n/a	0.5	0.8	1.8	4.1	7.1	10.0
Average selling price			n/a	n/a	n/a	$30.29	$38.00	$16.96	$15.29	$14.25	$28.90

*16K 3PS refers to the industry-standard, three-power-supply DRAM. The 16K 5V model requires only one power supply.
**Sales of 1K DRAMs were negligible by 1977. Estimates are created by assuming Intel prices at average selling price. Casewriter estimates that by 1984 Intel DRAM sales were closer to $100 million. Losses to gross income due to DRAMs in 1984 were estimated by the casewriter to be between $20 million and $30 million.
† Architecture refers to company who originated design, not to manufacturer. For example, while Intel's designs captured 33.5% and 59.1% of the 8- and 16-bit segments, Intel's actual unit sales of microprocessors accounted for only 14.5% of total market sales in 1984. Licensing agreements with other vendors account for the remainder. Next to Intel, NEC was the second largest unit shipper of microprocessors at 13.5% of total units. Motorola captured fifth place behind Zilog (8.9%) and AMD (7.4%) with 7.3% of microprocessor unit sales in 1984.
Source: Dataquest

It was just another kind of memory at the time, and people saw it as a research and development device. Today, the likelihood of someone killing an effort like this one is very high, because we require a well-defined application to a market from the outset. This is especially so because we are not lacking in opportunities. There is still a lot of evolution left in the current technology. If you consider the possibilities for reducing line width, you can see another 12 years of evolution along the same curve.

The Invention of the Microprocessor

Ted Hoff invented the microprocessor. Intel had been hired by the Japanese firm Busicom to design and build a set of chips for a number of different calculators. Busicom had envisioned a set of around 15 chips designed to perform advanced calculator functions. Hoff suggested building a simpler set of just a few general-purpose chips which could be programmed to carry out each of the calculators' instructions.

He was the architect of the chip set which Federico Faggin and a team of designers implemented. The set included four chips: a central processing unit (CPU) called the 4004, a read-only memory (ROM) with custom instructions for calculator operation, a random-access memory (RAM), and a shift register for input/output buffering. It took nearly a year to convince Busicom that the novel approach would work, but by early 1970, Intel signed a $60,000 contract which gave Busicom proprietary rights to the design. The CPU chip, 4004, was eventually called a microprocessor.

While Intel produced chips for Busicom which were successfully made into 100,000 calculators,[*] a debate within the company developed about whether Intel should try to renegotiate the rights to the chip design. Hoff believed that Intel could use the devices as a general-purpose solution in many applications ranging from cash registers to street lights, and he lobbied heavily within the company.

Eventually, Intel decided to offer reduced pricing to Busicom in exchange for noncalculator rights to the design. Ed Gelbach remembered the management decision: "Originally, I think we saw it as a way to sell more memories and we were willing to make the investment on that basis." Busicom, in financial trouble, readily agreed to the proposal.

[*] The casewriter noticed a Busicom calculator on Gordon Moore's desk. Moore also wore an Intel digital watch, which he called his $15 million watch, referring to Intel's ill-fated venture into the watch business. He said: "If anyone comes to me with an idea for a consumer product, all I have to do is look at my watch to get the answer. . . ."

The 4004 was introduced in 1971. It contained 2,300 MOS transistors and could execute 60,000 instructions per second. Its performance was not as good as custom-designed logic, but Intel believed there was a significant market for it. Early on, it became apparent that Intel would have to educate its customers in order to sell the 4004. As a result, Gelbach's group developed the first of Intel's development aides, which were programming tools for the customer. By 1973 revenues from design aides exceeded microprocessor sales.

In tandem with the 4-bit 4004, Intel developed an 8-bit microprocessor, the 8008, which was introduced in April 1972. The 8008 was designed with a computer terminal company in mind, but was rejected by the company because it was too slow and required 20 support chips for operation.

In the meantime, Intel's advancements in static and dynamic RAMs had provided a new process technology which promised increased transistor switching speed. Intel had created an *NMOS* process, which was applied to the 8008. In addition, much of the functionality of the support chips was integrated into the new microprocessor, the 8080. As a result of process technology, the 8080 could execute 290,000 instructions per second. In addition, the 8080 required only six support chips for operation.

The introduction of the 8080 in April 1974 heralded the beginning of a new age in computing. The market for microprocessors exploded as new uses were developed. Intel was one year ahead of Motorola's introduction of the 6800 and eventually took nearly the entire 8-bit market. Even though the 6800 used an architecture more familiar to programmers, Intel offered more effective development aids and support systems. Several integrated circuit companies were licensed to produce the 8080 so that customers were assured of a second source of supply. Ed Gelbach remembered the mid-1970s as the good old days. At an initial selling price of $360 per chip, Intel paid for the 8080 research and development in the first five months of shipments.

Motorola and Zilog* continued to apply pressure in the 8-bit microprocessor marketplace (see Exhibit 4). But Intel's 16-bit microprocessor, the 8086, again was first to market by about one year when it was introduced in June 1978. Intel management decided that upward

*Zilog had been formed as a start-up by three Intel design engineers. Andy Grove commented that the loss of those engineers set back Intel's microprocessor program by as much as one year.

compatibility would be a critical feature of the 16-bit chip. While the 8086 could operate software developed originally for the 8080, it employed a new architecture which required new software for full exploitation. An 8-bit *bus* version of the new architecture, the 8088, was also introduced. For two years, Intel did not meet its sales forecasts for the 8086 family as customers purchased only sample quantities and worked on a new generation of software. In the meantime, Motorola introduced its own 16-bit microprocessor, the 68000, and appeared to be gaining momentum in the field.[7]

Recognizing that the 68000 represented a critical threat that could lock Intel out of the 16-bit market and potentially the next generation as well, Intel created a task force to attack the 68000. The project was called operation CRUSH. The project leader said: "We set out to generate 100,000 sales leads and get that down to 10,000 qualified leads resulting in 2,000 design wins in 1980." SWAT teams of engineering, applications, and marketing people were mobilized to travel anywhere in the world whenever a design win was threatened.

The CRUSH campaign emphasized Intel's systems approach, and produced 2,500 design wins in the first year. The most notable win was IBM's decision to use the 8088 in their first personal computer in 1981. IBM planned an open-architecture personal computer, and Intel's 8086 family defined the software standard. Intel sales representatives knew they won the IBM account several months before it was made public when the IBM Boca Raton office started placing orders for Intel's ICE-88 development systems. In 1981, 13 percent of Intel's sales were to IBM.

The project to develop the next microprocessor generation began in 1978. The 80186 and 80286 were designed to be upwardly compatible with the 8086, and to offer increased integration, internal memory management, and advanced software protection (security) capability. The 80286 was designed to operate with as few as four support chips. The 286 team developed product features through extensive field interviews, and created a list of over 50 potential applications ranging from business systems to industrial automation. Ironically, the applications list did not include personal computers, which later became the single largest application.

The 80286 was the most ambitious design effort ever undertaken at Intel. The chip contained 130,000 transistors (versus 29,000 for the 8086). Intel's computerized design tools were stretched to their limit. Four separate computer systems had to be used just to store the design. Design verification (a tool which checks that mask design correctly reflects schematic design) took four days of continuous computer operation. Several crises arose throughout the development period.

The 286 logic design supervisor recalled:

> At least once a year we went through a crisis that made us wonder whether we would get there or not. One was the chip size crisis. At one point, it looked like the chip would be as big as 340 mils on a side. That was so big that people outside the design team would roll on the floor laughing. They kind of enjoyed our misery. Chip designers love to hear that someone else's chip is too big, but when it happens to you, it's really serious stuff.

The design team of 24 people worked feverishly for three years to develop the first prototype. That device was fabricated in 1982 at Fab 3 in Livermore but did not operate with high enough speed. Gradually, all the bugs were worked out, and only one hurdle remained: developing the methodology to test the chips as they came off the line. Production was ready to start making the 80286 six months before the testing procedure could be developed. Intel had to develop computer tools in order to design the tests. The chip was introduced in 1983, 18 months later than originally planned.

In the meantime, Motorola was gaining momentum. Dennis Carter, who worked on marketing the 80286, said:

> The 68000 came out after the 8086 and it was having some success in the marketplace, but we weren't particularly concerned because we knew the 186 and 286 were on the horizon. We believed we would announce the 286, and everyone would flock to our door. But when we introduced it, the world perceived the 286 not as a powerful monster machine, but as a slight continuation of the 8086. It also seemed that a lot of startups were using Motorola, and that was real scary, because that's one indication of where the future is going to be.

Project CHECKMATE paralleled the earlier project CRUSH in concept. CHECKMATE task force members gave a series of seminars 200 different times to 20,000 engineers around the world. Rather than emphasizing performance specifications which Motorola could also use to advantage, the seminar stressed features which had been included at the request of the marketplace in 1978, such as *virtual memory addressing* and *multitasking*. Carter recalled:

[7] Motorola won the Apple computer account with the 68000. The 68000 architecture remains Apple's standard.

> As a result, the design wins completely turned around. When we went into CHECKMATE, some market segments were three or four to one in favor of Motorola. By the time we finished, it had turned around the other way.

Synergies Between EPROMs and Microprocessors

No one foresaw that microprocessors would create a booming market for EPROMs. The original four-chip design for the 4004 was general purpose except for the ROM chip, which had to be customized (at the factory) for each application.

Although it was developed separately, the EPROM substituted for the ROM and provided two advantages: the designer of a custom product could develop and revise the ROM-resident microprocessor programs quickly, and smaller applications which could not afford the expense of a custom ROM could substitute off-the-shelf EPROMS. Ed Gelbach commented:

> It made sense to be able to reprogram the microprocessor instead of buying fixed ROMs for it. You could change your system overnight or every five minutes with EPROM.

Intel had a competitive advantage in the EPROM process, and retained a majority market share until the late 1970s. Competitors had trouble imitating Intel's "floating gate" process. *Ron Smith*, manager of Static/Logic Technology Development, said:

> If a device physicist were confronted with the EPROM out of the blue, he might be able to prove it won't work. The EPROM process has as much art as science in it, not only in the wafer fab, but in the packaging, testing, and reliability engineering.

In 1977, Intel introduced the 16K EPROM, 2716, which was compatible with any microprocessor system. All alone with the floating-gate process, Intel enjoyed a boom in EPROM sales for two years.

By 1981, the industry faced a cyclical downturn, and Intel's virtual monopoly on the EPROM market was challenged by several competitors, including the Japanese. The industry average selling prices for the 16K EPROM dropped by 75 percent in 1980. Intel management responded by accelerating the introduction of the 64K EPROM.

In the midst of a semiconductor recession, Intel decided to retrofit the brand new Fab 6 at Chandler, Arizona, with a new photolithography technology: *Stepper alignment*. Fab 6 had just come online and was idle (see Exhibit 5 for more detail on Intel facilities). The gamble was significant: "new process, new product, new plant, and new people." The 64K EPROM (2764) team met very aggressive yield goals, and Intel was again leading the world in EPROM sales. By mid 1981, Fab 6 had produced hundreds of thousands of 2764s, and output was doubling every quarter.

Technology Development

The 2764 had been used by Intel's Santa Clara Technology Development Group to develop stepper alignment. Steppers allowed smaller feature definition and smaller die size,[8] but the capital equipment was an order of magnitude more expensive than conventional projection aligners. Because of the trend towards more expensive equipment and the growing need for a new generation of equipment for each generation of product, Intel modified its traditional philosophy of developing processes on fabrication lines.

From early on, Intel had divided its technology development into the three groups which represented the three major process areas: EPROM, DRAM, and logic. Competition between the groups for scarce resources in the Santa Clara facility had led to the decision to separate the groups geographically. By 1984, the three separate technology development groups were in three cities: EPROMs in Santa Clara, California; microprocessors and SRAMs in Livermore, California;[9] and DRAMs in Aloha, Oregon. While development of each technology was independent, management insisted on equipment standardization. Periodically, the groups got together, pooled information on equipment options, and agreed to purchase the same equipment.

Gordon Moore commented that resource allocation did not necessarily parallel the market fortunes of the process families:

> Allocation of resources to the different technology development groups is centralized by Andy and me. We want to maintain commonality. Also, we are old semiconductor guys. Ideally, one of the groups starts a new technology and the others follow. But for stepper technology this was not true; they all did it simultaneously.

The three groups each developed a distinctive style and distinctive competences which related to their product responsibilities (see Exhibit 6). The Santa Clara group was responsible for the EPROM and *EEPROM*

[8] Smaller die size leads to higher yield and lower manufacturing costs. If die size is reduced by 25 percent, manufacturing costs are typically reduced by at least 25 percent.

[9] The Livermore site was also a production facility in 1984.

EXHIBIT 5 Intel Facilities in 1984

Intel's wafer foundries

Fab area	Location	Year first opened	Original water size	Current wafer size	Technology development	Primary production focus
Ex Fab 1	Mountain View,	Purchased 1986	1"	Closed		
Fab 1	CA	1977	3"	4"	EPROM	Small number of EPROMs
Fab 2	Santa Clara, CA	1971	4"	4"	No	Logic
Fab 3	Santa Clara, CA	1973	3"	4"	Logic, SRAM	Logic and SRAM (was DRAM)
Fab 4	Livermore, CA	1979	4"	4"	No	Microcontrollers and EPROM
Fab 5	Aloha, OR	1979	4"	4"	DRAM	Pilot and DRAM
Fab 6	Aloha, OR	1980	4"	4" and 6"	No	Logic, EPROM, Microcontrollers
Fab 7*	Chandler, AZ	1983	5"	6"	No	EPROM only
Fab 8	Albuquerque, NM	Scheduled 1985	6"	6"	No	EPROM
Fab 9	Jerusalem, Israel	Scheduled 1986	6" plan		No	Under construction
Fab 10**	Rio Rancho, NM	Held at shell				
Fab 11**	Rio Rancho, NM	Held at shell				

*First 6" fab area in world. Original 5" facility used DRAMs for shakeout. 1981–82 recession delayed production and allowed installation of 6" equipment. Process transfer to 6" wafers was unexpectedly difficult and took over one year.
**These fab areas could be loaded with facilities and equipment and started in about two years

Intel's other worldwide facilities (excluding 50 sales offices)

Location	Date started	Product focus	Operation
Penang, Malaysia	1972	Broad	Component assembly and test
Manila, Philippines	1974	Broad	Component assembly and test
Haifa, Israel	1974	Logic	Design center
Barbados, West Indies	1977	Broad	Components assembly
Tsukuba, Japan	1881	Logic	Design center
Las Piedras, Puerto Rico	1981	Systems, DRAMs	Systems assembly, component test
Singapore	1984	Systems	Systems assembly

Source: Intel Documents

(electrically erasable programmable read-only memory) products. They focused on the processing steps most critical to EPROMs, for example, the double polysilicon process used to create the floating gate. Similarly, the Livermore group concentrated on processes critical to logic devices.

The DRAM technology development group led the company in linewidth reduction. For example, the DRAM group was developing a 1-μm process while the logic group was developing a 1.5-μm process. Two key factors made DRAMs suitable as a technology driver: large demand for the latest DRAM generation (early high-volume manufacturing experience) and simplicity of integrating design and testing with process development.

Process specialization in all three technology areas limited the direct transferability of processing modules from one area to another, but DRAMs still provided a convenient vehicle for leading-edge process learning, and the DRAM group was highly regarded. Ron Whittier said:

> In 1984, the memory technology development group represented Intel's best corporate resource for process development. People like Sun Lin Chou are a scarce resource in a technology-driven company. Sun Lin's group understands and executes process development better than any other group at Intel.

Dean Toombs described the DRAM group as different from the others because of the relationship between design and process engineers:

> The DRAM designer is a specialist and more a device physicist than other designers. He focuses on the memory cell and has to understand where every electron in the structure is. There is more of a connection between the designer and the process engineer. The design and the process are developed together. In contrast, a logic designer is not as concerned with the details of a transistor's operation. The process is critical, but not as interactive with the design.

EXHIBIT 6 Technology Development Groups

	DRAM	EPROM	Logic/SRAM
Location	Aloha, OR	Santa Clara, CA	Livermore, CA
Product focus	Moderate, undertakes some basic research	Strong, little basic research; EPROM and EEPROM development.	Strong
Process/design interface	Design engineers highly specialized in DRAMs with device physics focus. Process and design development are highly interactive and in parallel.	Process and design less tightly coupled.	Process and design loosely coupled. Design engineers focus on circuit design. Process engineers focus on shrinking linewidth technologies.
Key distinctive technical competence	Thin dielectrics and pushing photolithography limits. Tend to lead Intel in geometry reduction. DRAMS are seen as technology driver. Currently the only group with a 1-micron technology	Problems specific to EPROM and EEPROM. Expertise in developing polysilicon and passivation processes. Also focused on pushing technology to 1 micron.	Processes to shrink existing products and increase yields. Currently developing new process for 386 microprocessor. Developing expertise in double layer metalization.
Number of personnel	120	120	120
1985 budget allocation*	$65 million	$65 million	$65 million
Other comments	DRAM technology development group considered by many to be the most competent group. Major effort in 1-meg DRAM development. Facility has low turnover.	Relatively high turnover to competing companies in Silicon Valley. Has successfully maintained Intel lead in EPROM technology.	Technology development takes place in facility used for production of logic products. Major project in developing 386 process.

*Case writer's estimate.

Intel Product Line and Situation in Late 1984

By the end of 1984, logic products (including microprocessors, *microcontrollers*, and peripherals) were the dominant source of Intel's revenue (see Exhibit 7). The company offered over 70 peripheral chips which worked in tandem with its microprocessor lines. The 80186 and 80286 were tremendously successful. In addition to the IBM PC business, Intel had locked up the IBM PC clone business with customers such as Compaq, who purchased microprocessors either from Intel or from one of its licensed second sources such as Advanced Micro Devices. The only serious 16-bit architectural competitor was Motorola,[10] although *Electronic News* had reported that 10 companies, including NEC, Hitachi, Mitsubishi, Fujitsu, and Zilog, were developing proprietary 32-bit products, and National Semiconductor had already introduced its 32-bit offering. NEC's proprietary design effort was particularly interesting since NEC also supported Intel's microprocessor line as a second source.[11]

Intel had also developed a line of microcontrollers which integrated logic and memory (both SRAM and EPROM) to provide a self-sufficient, one-chip computer. One Intel manager suggested that integration of EPROM technology with logic was an effort to lift EPROMs from a commodity status. The microcontroller business had products in the 4-, 8-, and 16-bit market segments, which were used to control everything from house fans to complex satellites and had prices ranging from one thousand to several thousand dollars per chip.

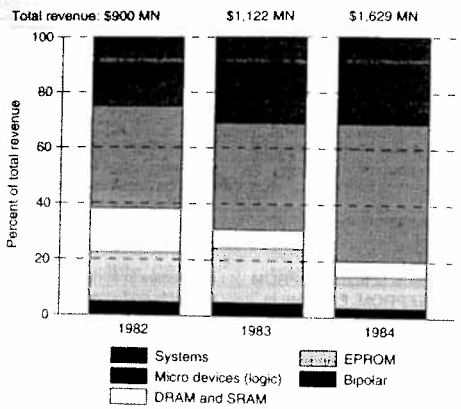
Scheduled for introduction in late 1985 was the successor to the 286, the 32-bit 80386™ microprocessor.[12] According to one Intel manager, "Once again, Intel was

[10] Motorola's 68000 has a 16-bit bus but actually uses a 32-bit internal architecture.

[11] *Electronic News*, February 18, 1985. The article also reports that Fujitsu did not confirm rumors that it had a proprietary 32-bit design. Instead, Fujitsu indicated its development efforts were still centered on second source agreements with Intel.

[12] 386 is a trademark of the Intel Corporation.

EXHIBIT 7 Composition of Revenues



Source: Dataquest.

betting the company on a new product." With 270,000 transistors, the 386 was even more complex than the 286. Intel had invested heavily in computerized design and simulation tools which made the design task run more smoothly. In 1984, Intel believed it had the best chip design capability in the world. However, Motorola had developed a strong 32-bit product, the 68020,[13] and was already in the marketplace winning designs, locking customers into its architecture.

The 80386 was scheduled to be one of the first products made with the new *complementary MOS (CMOS)* process (the 80C51 microcontroller and the 51C64 DRAM had both used versions of CMOS). It was also the first microprocessor to use stepper alignment; *double metalization*, and *plasma etching*. Development of the 386™ process was taking place in parallel with a new SRAM process at Livermore under the direction of Ron Smith. Ron Smith explained:

> Our group was called the Static Logic Technology Development Group and our charter was to develop *scaling improvements*[14] for the logic and SRAM lines. SRAMs were to lead the company in scaling. We saw the SRAMs not only as a product line but as a vehicle for microprocessor development. The SRAM is an indispensable tool in developing any new process. It is much easier to debug a process using memory components, because they are easier to test. That's why Intel traditionally uses memory products to develop a new technology.

In 1984, the Livermore group was developing two distinct processes, since the performance requirements for SRAMs and microprocessors differed. Although Intel had a good position in the low-volume, high-speed SRAM segment, it did not participate in the largest SRAM segment, which demanded higher density (more storage capacity).[15] The high-volume SRAM segment demanded a new four-transistor cell design and process. By contrast, the high-speed SRAM and the new 80386™ microprocessor both demanded a six-transistor CMOS design.

The high-volume SRAM process required a complex *polysilicon resistor* technology which was giving Smith's group difficulty. Smith described the environment as it had evolved in mid-1985:

> Eventually, we decided to drop the poly resistor process and go with a six-transistor CMOS SRAM product so that we could focus our attention on the 386 development. Basically, we sacrificed the high-volume SRAM for the 386™.
>
> To get an idea of the complexity of the 386™ development, compare it to the 286. The 286 team really comprised only six people. When it came time to develop the 386™, we had to come up with a double metalization process while at the same time reducing line widths to 1.5μm (from 2μm) and implementing the CMOS process. The 386™ process team had about 60 people: specialists in plasma etching, stepper alignment, chemical etching, and diffusion. If you compare the mask design for the 286 with the 386, you'll be able to tell how much area we saved by going to a double-layer metal. Lots of the 286 area was taken up with the routing of metal.

Gordon Moore described a linkage between market and technology development which may have contributed to the loss of a competitive SRAM product.

> Product designers want to see their product in high volume. So, it is important to have volume in a product line to get high-quality designers on board. For instance, SRAMs received less attention for that reason than I wish they had. We had a strong position in high-speed SRAMs, but we gave it up without really making a conscious decision.

The systems business at Intel had continued to grow with the company and by the end of 1984 represented the same 30 percent of revenue that MSO had represented in 1973. While a great deal of the systems business comprised development products aimed at microprocessor and microcontroller users, Intel also had vertically integrated into software development systems and single-board computers so that it could offer its customers options at several levels of integration.

[13] Introduced in sample quantities in September 1984.

[14] Scaling improvements allowed Intel to reduce the chip size of existing products without expensive redesign. The reduced chip size led to reduced manufacturing costs.

[15] Intel's overall SRAM position had diminished significantly over the years as Japanese manufacturers gained market share.

Manufacturing and Process Fungibility

While tolerating some process proliferation within plants, Intel took great pains to standardize each facility as it expanded its manufacturing base. In 1973, Grove was pictured in a snapshot at his desk with a foot-long mock chip package. On its side was printed the McDonald's Golden Arches logo with "McIntel" substituted. Each Intel chip would "look and taste the same no matter which facility produced it."

As larger-diameter silicon wafers became available, Intel developed a process on one line and then transferred the technology to its other facilities. For example, a process for 4-inch-diameter wafers was first developed at Fab 3 in Livermore, California, by a team of three people. The team leader then supervised the start-up of Fab 5 in Aloha, Oregon, which was dedicated to 4-inch wafers. In 1983, after delaying start-up due to the 1981–82 recession, Intel was the first semiconductor company to use 6-inch wafers at Fab 7, Rio Rancho, New Mexico (see Exhibit 8).

By 1984, Intel had seven fab areas in the United States, all within a two-hour flight of headquarters in Santa Clara. Due to more stringent manufacturing standards, the cost of a fab area had risen dramatically since the 1970s. A new fab area fully equipped cost between $150 million and $200 million and took about two years to construct. The first overseas fab area had just opened in Jerusalem, Israel. *Jack Carsten*, senior vice president and general manager of the Components Group, commented in retrospect on the decision to locate in Israel:

> Around the time we were deciding to put up a fab in Israel, I supported the idea of building a fab area in Japan. I had actually obtained leases on Japanese soil so that Intel could locate its first overseas fab area in Japan. That plant would have provided some insulation from currency fluctuations, but the Israel plant had tremendous government subsidies and a good labor market. A Japanese plant would have also put us into the pipeline of Japanese equipment vendors, and linked us into the Zaibatsu network. We could have tapped the expertise of Japanese DRAM technology development, silicon makers, mask makers, and the infrastructural support. This is what Texas Instruments did, because they had a commitment to local manufacturing. Eventually, we chose Jerusalem, largely because of the subsidies. This is not to say that the Israel facility is bad. It is a fine facility, but it certainly can't offer currency hedging against the Japanese yen.

Nearly all (97 percent) manufacturing capacity was devoted to MOS devices. Within MOS, the majority of processing was NMOS, but there was a trend towards increased CMOS. Each production facility was more or less dedicated to a particular process family (DRAM, Logic, or EPROM), although some facilities manufactured more than one family. Within each family, some process sequences were sometimes customized to accommodate particular product performance needs. While the equipment within any fab area was similar, different fab areas had different generations of equipment, and some processes required more of a particular machine for line balancing. Gordon Moore commented on the proliferation of process technologies:

> Over time, there has been a tendency to get more and more processes, and that complicates manufacturing allocations. In the past, we solved the problem by brutally getting out of businesses. But the customers didn't like that. For instance, we abdicated share in microcontrollers because we had to clean out somewhere to do other things.

While each facility could not produce every family of products, there was some fungibility between products and facilities. In times when demand was strong and capacity constrained sales, Intel division managers would get together monthly to decide how to load the factories. The chief financial officer, *Bob Reed*, described the process as being one that maximized margin per manufacturing activity:

> Basically, there are three main process areas: fabrication, assembly, and test. Assembly is usually not a constraining factor—you can ramp it up as fast as you need to. Similarly, test can be ramped up in the short term. Fabrication (the front end of the process) is usually the bottleneck in times of tight capacity—it takes long lead times to increase capacity. Since fabrication is the constraining resource, fabrication is the key variable for assigning cost to products.
>
> Each process sequence (EPROM, Logic, or DRAM) is assigned a total amount of manufacturing activity based on the number of steps it requires. Total company manufacturing costs are then allocated to products on the basis of manufacturing activity. For each product, the overall yield (number of good die at final test versus total number of die on starting wafer) is applied as a divisor to the process cost to arrive at a total cost per good part. The sales price per part

EXHIBIT 8 Sample of Cost Accounting Data for Selected Intel Products in 1984

Product	Process	Raw wafer cost	Number of mask layers	Number of activities	Cost per activity	Line yield	Cost per wafer	Die per 6" wafer	Wafer sort yield	Total cost per die	Package/test cost per die	Yield at test	Total cost per chip	Average selling price	Contribution margin per chip
64K DRAM	NMOS DRAM	60	8	30	72.00	90%	2,467	1900	90%	1.44	0.45	90%	2.103	2.05	-2%
64K DRAM	CMOS DRAM	100	10	38	72.00	84%	3,376	1806	85%	2.20	0.45	90%	2.944	3.08	4%
256K DRAM	CMOS DRAM	100	10	38	72.00	83%	3,417	922	60%	6.18	0.65	90%	7.585	16.27	53%
64K EPROM	NMOS EPROM	60	12	48	72.00	79%	4,451	1582	75%	3.75	2.65	90%	7.112	8.15	13%
256K EPROM	NMOS EPROM	60	12	48	72.00	78%	4,508	756	60%	9.94	2.45	90%	13.764	21.00	34%
80286	LOGIC	60	10	40	72.00	90%	3,267	172	70%	27.13	2.00	85%	34.273	250.00	86%
80386 (samples)	1.5 μm LOGIC	100	13	50	72.00	90%	4,111	131	30%	104.61	15.00	85%	140.716	900.00	84%

Key:

Raw wafer cost: raw wafer cost differs depending on whether or not process is CMOS.
Number of mask layers: Refers to the number of times the wafer goes through the photolithography step.
Number of activities: Basic unit of manufacturing for cost accounting purposes. Refers to the number of times the wafer is physically altered in the process.
Cost per activity: An average of worldwide manufacturing costs, including depreciation, materials, labor, and other facilities costs.
Line yield: Ratio of wafers started to wafers completed.
Die per 6" wafer: Number of devices on a 6" wafer (function of die size)
Wafer sort yield: Number of good die divided by total die after all processing is completed and before wafer is sawed and devices are packaged.
Total cost per die: Cost per wafer divided by number of good die per wafer at wafer sort test.
Packages/test cost: Cost of packaging and testing one device.
Yield at test: Number of devices entering packaging divided by number of devices which pass final test.
Total cost per chip: Total cost per die plus packaging and testing costs all divided by yield at test.

Source: Casewriter estimates.

is then used to calculate margin per part, and margin per activity (see Exhibit 8).

According to Reed, sometimes the numbers told a compelling story about the DRAM business. The difference between margin/activity for DRAMs and for the highest margin products could be an order of magnitude. Ron Whittier, general manager of the Memory Components Division from 1975 until 1983, felt that the system for plant allocation was a very good one:

> Some companies really went too far by selling capacity to the highest bidder within the company. At Intel, a minimum production allocation would be assigned based on how much we needed to produce to maintain our long term market position. Basically, we used our independent distributors as buffers. In times when DRAM production was pressured by other products, we tapered sales to independent distributors while maintaining sales to large account customers.

Grove commented that since the distributors never accounted for more than 20 to 30 percent of Intel's DRAM business, they could not really account for the leveling in Intel's DRAM sales (see Exhibit 4).

Whittier also noted that DRAMs had at one time been the single largest product line and thus could not easily be entirely displaced by other products unless total capacity was decreased. The finance group thought of DRAMs as a "low ROI, high beta" product line. Bob Reed insisted that the DRAM manager sign a symbolic check equal to the margin foregone whenever high-margin products were bumped by DRAMs.

Ed Gelbach explained why Intel had stayed with the DRAM even though it looked less profitable than other products:

> I was in favor of keeping DRAMs from a marketing strategy standpoint. A full-line supplier has a basic advantage in any sales situation. When you're competing with full-line suppliers, it helps to be able to offer a comparable line. Since customers often pay particular attention to their highest-dollar-volume vendor, it also pays to offer the commodity product since it is generally purchased in high volume. A more subtle reason boils down to reputation. Intel had been known to drop unprofitable products, sometimes leaving customers high and dry.
>
> In board meetings, the question of DRAMs would often come up. I would support them from a market perspective, and Gordon [Moore] would support them because they were our technology driver. Andy [Grove] kept quiet on the subject. Even though it wasn't profitable, the board agreed to stay in it on the face of our arguments.

ENVIRONMENTAL FORCES

Bob Reed realized the entire U.S. semiconductor industry was in trouble even during the boom year of 1984.

> Even though ROS for the industry was relatively high in 1984, asset turns were decreasing and ROA was low. The business had become too capital intensive. An astute observer could see that the U.S. industry as configured couldn't provide its investors with an adequate return when a new plant cost $150 million and took at least two years to build. Intel was virtually alone with a respectable ROE.

In 1985, the semiconductor industry was expected to enter into another in a series of cyclical down turns which seemed to occur every five years. The cause of the cyclical recessions was a classic case of oversupply and softening demand. Since 1980, a large amount of worldwide semiconductor fabrication capacity had been added, and the learning curve effect (increase in yields, decrease in chip size, etc.) added another 30 percent per year to worldwide capacity.

In the previous recession, Intel had been one of a few companies not to cut back its production workforce. While Intel did not have a no-layoff policy, during the 1981–82 recession Andy Grove had instituted the "125 percent solution." In that program all salaried employees were asked to work an additional 10 hours per week without additional compensation to accelerate product introductions. When the 1980 recession proved to be longer than expected, Intel instituted a 10 percent pay cut in addition to the 125 percent solution.

Intel had several groups of competitors (see Exhibit 9). The first were other U.S. full-line digital design and supply houses such as Motorola, National Semiconductor, and Texas Instruments (TI). Motorola had made the transition from a tube manufacturer in the 1950s to a diversified semiconductor and electronic systems manufacturer in the 1980s. It offered a full line of products competitive with Intel's, including DRAMs, microcontrollers, and microprocessors and was Intel's only serious challenger in microprocessor architecture. TI, while not renowned for its microprocessors, also had a complete product line, including a facility in Japan which was fabricating DRAMs.

The second category of competitor focused on process technology as opposed to design. That group was represented by AMD. While AMD produced a full line of component products, a significant portion was manufactured under license from Intel and others.

The third group included foreign competition, particularly Japanese. Japanese competitors included Hitachi,

EXHIBIT 9 Selected Competitor Data for 1984

FY 1984 (in millions of dollars)	Intel	National Semi-conductor	Texas Instruments	Advanced Micro Devices	Motorola	Hitachi	Toshiba	NEC	Fujitsu
Semiconductor sales	$1,201	$1,213	$2,484	$ 515	$2,319	$2,051	$1,516	$2,251	$1,190
Total sales	1,629	1,655	5,741	583	5,534	18,528	11,003	7,476	5,401
COGS	883	1,146	4,190	276	3,206	13,632	8,182	5,117	3,346
R&D	180	158	367	101	411	898	597	391	(incl.)
SG&A	315	247	491	108	1,064	3,367	2,758	1,443	1,453
Other	(48)	1	168		387		(1,106)	673	335
Profit	299	103	525	98	466	631	572	367	523
Profit after tax	198	64	316	71	387	709	250	189	297
Depreciation	113	115	422	43	353		627		374
Capital expenditure	388	278	705	129	783		1,192	883	747
Total assets	2,029	1,156	3,423	512	4,194	7,997			5,699
LT debt	146	24	380	27	531	1,379	1,830	1,524	915
Total equity	1,360	619	1,540	278	2,278	6,118	2,191	1,728	1,935

Semiconductor market share in 1984	Bipolar Digital	EPROM	DRAM and SRAM	MOS Micro-component	MOS logic	Linear	Discrete	Opto-electronic	Total (in millions)
AMD	5.4%	10.5%	0.5%	1.8%	0.1%	0.4%			$515
Fairchild	8.6%		0.1%	0.5%	0.7%	2.9%	1.3%	0.3%	665
Fujitsu	6.4%	11.1%	7.8%	3.7%	3.0%	0.8%	0.8%	4.3%	1,190
Hitachi	4.7%	17.4%	15.1%	3.7%	2.2%	3.7%	8.6%	4.3%	2,051
Intel	0.7%	16.0%	3.4%	23.0%	1.2%				1,201
Mitsubishi	2.6%	13.3%	4.0%	4.8%	0.4%	2.1%	3.7%	1.1%	964
Mostek			7.1%	1.7%	1.8%				467
Motorola	9.5%	1.1%	6.1%	9.0%	10.4%	5.5%	12.2%	1.6%	2,319
National	6.1%	4.2%	1.1%	3.6%	5.9%	8.9%	0.9%	1.2%	1,213
NEC	2.6%	5.8%	13.0%	12.7%	8.3%	5.9%	7.6%	2.8%	2,251
Philips	12.3%		0.7%	3.2%	3.7%	4.8%	4.4%	1.5%	1,325
TI	22.5%	10.5%	10.8%	3.6%	2.9%	8.4%	1.2%	4.1%	2,484
Toshiba	0.8%	3.6%	7.1%	2.2%	8.7%	4.7%	8.4%	8.8%	1,516
Others	14.4%	6.5%	21.3%	21.9%	50.3%	51.8%	50.9%	70.1%	10,900
Total market (in millions of dollars)	$4,783	$1,319	$4,906	$3,229	$3,493	$4,888	$4,986	$1,221	$29,061

Key:
MOS microcomponent: microprocessors, peripherals, and microcontrollers.
MOS logic: gate arrays, custom logic, and application-specific ICs.
Linear: operational amplifiers, comparators, and other analog devices.
Discrete: single transistors, diodes, and thyristors.
Optoelectronic: LEDs, semiconductor lasers, and solar cells.

Source: Dataquest and annual reports.

Fujitsu, NEC, Toshiba, and others. They had concentrated primarily on DRAM and SRAM products, although each also had a significant share of the EPROM market and served as second sources to U.S. microprocessor and microcontroller suppliers. Intel had second-source agreements for its microprocessor line with Fujitsu and NEC.

Several U.S. DRAM makers had accused Japanese manufacturers of dumping DRAMs at prices below cost throughout the early 1980s.

Industry observers saw that Japanese firms under the direction of MITI had targeted semiconductors as a strategic industry and were investing for the long term. In the years between 1980 and 1984, U.S. firms invested a total of 22 percent of sales in new plant and equipment while Japanese firms invested 40 percent. The result was that by 1983, Japanese total investment in semiconductors exceeded U.S. investment. Production yields of Japanese semiconductor companies exceeded those of U.S. producers by as much as 40%.[16]

DRAMs were not the only product under siege by the Japanese. *The Wall Street Journal* published a story in June 1984 which reported on a memo sent by Hitachi to its U.S. EPROM distributors. The memo said: "Quote 10% below their price; if they requote, go 10% again, don't quit until you win."[17]

Intel had been wary of Japanese semiconductor companies for some time and had sued NEC in 1982 when it alleged NEC copied its 8086 product without license. Peter Stoll, an 8086 designer at Intel, realized his chip had been copied when he discovered that NEC's 16-bit microprocessor had two transistors which were disconnected from the rest of the circuit at exactly the same place where he had disconnected them in a late revision of the Intel mask set.[18] This was considered evidence that NEC copied the chip without even understanding its design.

Bob Reed emphasized the importance of Intel's ability to protect its intellectual property:

> If our primary value added is in our design capability, we've got to protect that with vigilance. We have a strict policy of pursuing anyone or any company that appropriates our intellectual property—design or process.

In this highly competitive environment, managers at Intel and other companies often had to consider the problem of spin-off companies. Key engineers had sometimes left Intel to form their own companies with venture capital help. Their departure would stall research at a minimum and, according to Gordon Moore, could be seen as diluting the U.S. industry's ability to compete. Spin-offs were sometimes accused of taking technology with them.[19]

DRAM SITUATION IN 1984

Loss of Leadership Position

By the end of 1984, Intel had lost significant market share in DRAMs (see Exhibits 4 and 9). The first real difficulties had come with the 64K generation. In 1980, Intel's 5-volt 16K DRAM was still a market success due to process innovations, and work was continuing on the 64K generation. DRAMs traditionally led the company in new technology development, and the 64K DRAM was no exception.

Ron Whittier said that to make the 64K version, the memory cell size was reduced, but the actual die size still had to be increased significantly. The DRAM group calculated that given current defect levels in manufacturing, the required die size would be too big. Based on the number of defects per square centimeter normally experienced in fabrication, the projected yield on the 64K DRAM would be too low to be acceptable. In order to boost yield, the group decided to build in redundancy at the chip level.

Whittier described the redundancy technology:

> Essentially, you have a row-and-column addressing system on a memory chip. The periphery of the chip contains logic and refresh circuitry necessary to control and update the DRAM. In the 64K version, Intel added an extra column of memory elements so that in the event of a process-induced defect, the auxiliary column could be activated. There was a physical switch, or "fuse," built in to each column which could be addressed by the tester machinery. When a bad element was detected, current would be passed through the switch and would blow a "fuse," inactivating the defective column and kicking in the auxiliary column. In this fashion,

---

[16] Clyde Prestowitz, "While the best U.S. companies obtained yields of 50–60 percent, the best Japanese were getting 80–90 percent," *Trading Places*, 1988, p. 46.

[17] *The Wall Street Journal*, June 5, 1985.

[18] Clyde Prestowitz, *Trading Places*, 1988, p. 48.

[19] Intel had sued SEEQ for taking a technology for electrically erasable PROMS (EEPROMs). Excel, a spin-off from SEEQ was later sued by SEEQ. Note: Intel continued to pursue its own EEPROM process but eventually decided not to participate in that market because it was too small. A second engineering team left Intel on friendly terms to found Xicor. In 1985, Xicor and Intel were negotiating a joint research project.

a defective memory chip could be "reprogrammed" before shipment, and overall yield could be improved.

Dean Toombs, general manager of the memory components division after 1983, had worked on DRAMs at Texas Instruments (TI) before coming to Intel. Toombs said the discussion on redundancy was industrywide. At TI, engineers had concluded that at the 64K generation redundancy would not be economical and had deferred the discussion until the next generation. For the 64K generation, TI ultimately chose to focus on reducing the defect level in manufacturing.

Intel's redundancy program started out successfully. Two 64K DRAM projects were carried out in tandem, one nonredundant and the other redundant. Prior to production commitment, the redundant design was a clear winner, with yields over twice that of the nonredundant design.

Success quickly turned to failure as a subtle but fatal defect in the redundant technology showed up late in development. The fuse technology was less than perfect. The polysilicon fuse would blow during testing as designed, but a mysterious regrowth phenomenon was detected during accelerated aging tests. Sun Lin Chou commented:

> The failing-fuse problem was simply a case of not having done enough engineering early on. We just didn't fully characterize the process technology and the fusing mechanism.

The result was that the switch eliminating the defective column of memory cells was not permanent. In some cases, the device would revert to its original configuration after being in the field for some time—meaning the defective cell would again become a part of the memory. Errors would occur in which the device alternated randomly between the two states, meaning that at any given time the location of data stored in the memory became uncertain. In either case, the failures were not acceptable, and Intel could not develop a quick fix.

In the meantime, Japanese competitors were throwing capacity at 64K DRAMs and improving the underlying defect density problem which Intel's redundancy program had meant to address. Between July 1981 and August 1982, Japanese capacity for 64K DRAM production increased from 9 million to 66 million devices per year.[20] Whittier took a one-week trip to see Intel sales engineers[21] and explain that Intel's 64K DRAM would be late:

> The sales force was very disappointed in the company's performance. Any sales force wants a commodity line. It's an easy sell and sometimes it's a big sell. That trip was perhaps the most difficult time in my whole career. When I announced we would be late with the product, the implication was that Intel would not be a factor in the 64K generation.

While the development team eventually fixed the fuse problem and was the first to introduce a redundant 64K DRAM, the 2164, its introduction was too late to achieve significant market penetration.

Attempts to Regain Leadership Position

Having assessed that they were behind in the 64K DRAM product generation, the DRAM group took another gamble. The development effort was shifted from NMOS to CMOS. The advantage of CMOS circuitry was lower power consumption and faster access time. Intel defined a set of targeted applications for the CMOS DRAM technology.[22] Whittier's strategy was to introduce the CMOS 64K and 256K DRAMs in 1984. The notion was that by creating a niche market with premium pricing, Intel could maintain a presence in the DRAM market while accelerating forward into a leadership position at the 1-meg generation.

Dean Toombs said that by the time he took over the Memory Components Division in 1983, things were "clicking along." Demand was in an upswing, and Intel seemed to have a technology strategy which could lead to dominance in the 1-meg DRAM market. Many of the 2164 sales in 1983 went to IBM, and in addition Intel sold IBM the 2164 production and design technology. Toombs recalled that in late 1983 and early 1984, the silicon cycle was on an upswing and memory product demand was at an all-time high. The memory components division's bookings exceeded its billings.

During the boom of late 1983 and early 1984, all of Intel's factories were running at capacity. Allocation of production capacity between products was necessary. The question facing the memory components division was how to effect the transition from NMOS to CMOS.

[20] Clyde Prestowitz, *Trading Places,* 1988, p. 44.

[21] Intel sales engineers sold Intel's entire product line but were supported by applications engineers in a ratio of one engineer to every two sales representatives.

[22] One such application was laptop computers, which place a premium on low-power consumption chips.

Toombs said the "hard decision" was made to completely phase out the NMOS line. All DRAM fabrication was consolidated in Oregon's Fab 5. Toombs suggested that the decision to "go CMOS" was consistent with Intel's general philosophy: to exploit new technology and create a lead against competitors based on proprietary knowledge.

The development of the CMOS 64K and 256K DRAMs took place in a facility adjacent to the Oregon production facility. While the development was not on the production line, there was a fairly smooth transition into manufacturing. The CMOS technology was more complex, requiring 11 to 12 masking steps versus 8 to 9 steps for NMOS. This resulted in a higher manufacturing cost for the CMOS process (see Exhibit 8).

The CMOS DRAM products were introduced in 1984 and priced at about one and a half to two times the prevailing NMOS price. Intel management developed a niche strategy: differentiate the product from other offerings and sell it on features. In addition to the CMOS feature, Intel offered an alternative memory organization which provided performance advantages in some applications. Intel sampled the products broadly to many customers and made many design wins, particularly in situations where other DRAMs had inadequate performance. The 256K chip was well-designed and executed. Sun Lin Chou commented:

> The 256K CMOS DRAM was the first DRAM product which did not have to go through some sort of design or process revision before or after going to market. With this product, we felt we were regaining our lead in DRAM technology after three generations.

The CMOS DRAMs started as a winning product family. Unfortunately, the market softened as 1984 went along. The price of NMOS DRAMs fell by 40 percent in one three-month period from May to August 1984. In the scramble and upheaval of the semiconductor market, Toombs said that Intel's differentiation message got lost. All suppliers were pushing products into the market, and Intel's superior product specifications seemed like just another ploy to get volume.

By late 1984, Intel's ability to make profits and, more importantly, to project future profits in DRAMs was limited. Said Toombs: "In a commodity marketplace, your staying power is a function of the size of your manufacturing base." According to Toombs, by late 1984, Intel was down to less than 4 percent of the 256K DRAM market and had lost its position entirely in 64K DRAMs.

On the other hand, the technical strategy seemed to work, since the first prototype of the 1-meg DRAM was expected in March 1985. However, as Sun Lin Chou indicated, Intel's technology strategy for the 1-meg DRAM had been different from that of previous generations:

> Our advanced capability in thin dielectric has allowed us to focus on reducing the minimum feature size to one micron instead of changing the entire cell design. Some memory leaders have chosen to scrap the traditional capacitor design, and are trying to move to a smaller "trench" capacitor which requires an entirely new generation of equipment and processing. While they are still at 1.2 to 1.5 microns, we've pushed the photolithography technology further. We may have to go to the trench capacitor in the next generation [4 megabit], but by then we will be able to take advantage of their learning.

Toombs believed that the DRAM technology development group had provided Intel with a unique product capability:

> The 1-meg DRAM will be a technically outstanding product, at least one and a half to two years ahead of any competition in application of CMOS. But the handwriting is on the wall. In order to make the DRAM business go, major capital investment is required and the payback just isn't there. The issue for 1985 is how to survive.

Jack Carsten believed it was critical for Intel to stay in the DRAM business. But in case the company was no longer willing to dedicate facilities to DRAMs, he felt a technology transfer deal should be made with a Korean chip manufacturer:

> The play I am proposing is to stop manufacture of the DRAMs, and to form an alliance with a large Korean company who has state-of-the-art capacity installed. We now have a functional 1-meg DRAM. Basically, Intel could support the business through an R&D alliance and be the technology leader.
>
> To be fair, you have to realize that the Koreans have state-of-the-art equipment, but are not yet expert at using it.[23] In order to make the technology transfer work, we would have to transfer 20 or so of our crack engineers to teach the Koreans how to make the 1-meg DRAM. Apart

[23] Note: In February 1985, Intel was to enter into an agreement with a Korean firm to transfer technology for two Intel parts. The technology had been developed at Intel to introduce the 8048 microcontroller (same generation as the 8085 microprocessor) and the 2764 EPROM (see Exhibit 3 for timeline). While those processes required 3- to 4-micron geometries, the 1-meg DRAM product required 1-micron geometries. The Korean company had annual semiconductor sales of about $10 million in 1984.

from the technology risk, there is the risk that we would create a new competitor. History is rife with examples of how technology transfers have backfired, and we've certainly been burned before. But, maybe there's some truth to the logic that the enemy of your enemy is your friend.

OPTIONS FOR DRAM

Grove could see several distinct options for the DRAM business: (1) drop it all together, (2) stay in the business as a niche player, (3) license the technology to another company, or (4) invest in DRAM capability at the 1-meg level and commit to a low-margin business.

As he reflected on the situation, he thought about how Intel had arrived at its current position:

> At the 16K level, we were leading in both EPROM and DRAM products, but capacity was tight. We reduced our commitment to DRAMs in what was, in effect, a capital appropriations decision. Margins and customer dependence were both important in causing us to shift our focus to EPROMs.
>
> Then came the lackluster 64K design. We stumbled and it was a burning embarrassment. Our market position was at 2 to 3 percent. You just can't win like that.
>
> Gordon [Moore] is probably right when he says the only difference between DRAMs and EPROMs is that EPROMs never missed a turn. If you miss a turn, the game is over.
>
> The bright side is that we might have lost a lot more if our 64K generation had been a success. Texas Instruments is probably losing more than five times what we are.
>
> We have been trying to find a clever way to stay in this business without betting everything we have, but maybe there is none.
>
> The key question is, Should we really commit to being a leader? Can we be? What is the cost if we try? What is the cost if we don't?

TECHNICAL APPENDIX

Access: In this context, refers to the circuitry which allows the DRAM user to read and write to specific locations of memory. Access time is a critical performance feature of DRAMs and refers to the amount of time it takes to read or write a bit of memory. Often DRAMs offer two different access modes, one that is bit by bit and one that writes or reads large amounts of data. The bit-by-bit rate is typically slower.

Bipolar: Refers to a generic type of transistor and to the family of processes used to make it. The bipolar transistor consumes more power than the MOS transistor but can be made to switch faster. Excessive power consumption limits the density of bipolar products. The bipolar process is a relatively complex semiconductor process.

Bus: Refers to the communication backbone of the microprocessor. An 8-bit bus can transfer 8 bits of data at a time between the microprocessor and the outside world (memory or other peripherals). The 8-bit-bus version of the 8086 actually has a 16-bit internal bus. Each cycle within the chip can handle two cycles of data input.

Capacitor: A circuit element (transistors, resistors, capacitors) which consists of two metallike layers separated by a thin insulating film. In a typical integrated circuit the silicon substrate (wafer) acts as the first metallike layer. The silicon surface is oxidized to form the insulating layer (silicon dioxide) and then a polysilicon layer is deposited over the oxide to form the second metallike layer. In the context of DRAMs, the capacitor acts as an information storage device. When a positive charge is placed on one surface of a capacitor, a negative charge is induced on the opposite surface. The capacitor holds the charge for a limited period of time, and the presence of the charge indicates a bit (binary digit) of information. The ability of the capacitor to store charge is related to its area and the thickness of the insulating film. The thinner the insulator and the larger the surface area, the more charge a capacitor can store. (See *trench etched capacitor* for more information.)

Chip: Refers to the actual integrated circuit which is cut from the wafer after fabrication. Typical chips are 100–400 mils on a side and can contain several hundred thousand transistors. The chip is put into a package where microscopic wires are attached to the die and brought out of the package in larger pins which can be soldered into a printed circuit board.

Class 10 production facility: Semiconductor fabrication plants are perhaps the cleanest areas ever created. Airborne particulates such as dandruff, pollen, and other forms of dust are a major source of semiconductor manufacturing yield problems. One particle of dust settled on a silicon wafer is enough to ruin an entire chip. The class number of a facility refers to the amount of particulate in the air. Class X means that 1 cubic foot of air on average will contain X or fewer particles. A class 10 fabrication facility is designed with advanced air-filtering designed to eliminate turbulence. Operators wear specialized clothing and enter clean rooms only through air showers which remove contamination. To give a sense of the cleanliness, a typical hospital operating room is between class 1,000 and 10,000.

Complementary MOS (CMOS): Refers to a semiconductor process which can produce a specific configuration of transistors which include both NMOS and PMOS devices. A group of six transistors fabricated in CMOS forms the fundamental building block for Intel's latest generation of logic circuitry. The six-transistor cell is a bistable cell which is either in the on or off state. CMOS has the advantage of very low power consumption, since none of the transistors ever draws current except during the time when the six-transistor cell changes states from on to off. Laptop computers use exclusively CMOS integrated circuits.

Die: See *chip*.

Dielectrics: Refers to insulating materials. In semiconductor processing they include silicon dioxide, silicon nitride, silicon oxynitride, and others. Dielectrics are used in several areas of integrated circuits. In DRAMs, they are used for storage capacitors. In MOS transistors, they form the gate insulator.

Double metalization: Until the 80386, all of Intel's circuits employed only one layer of metalization. The design of logic circuitry (where interconnection between groups of transistors appears to be random) is greatly simplified by adding a second layer of metal. Although the processing sequence is complicated, double-layer metalization allows chip size to be reduced.

Dynamic random-access memory (DRAM): A variety of RAM which maximizes utilization of silicon "real estate" and minimizes power consumption per storage bit. Each bit of information is stored as a charge on a capacitor driven by one transistor. Since the charge dissipates rapidly even when power is constantly supplied to the device, the information within each memory location must be rewritten (refreshed) hundreds of times a second. While the refresh function was originally taken care of by external circuitry, the latest DRAM chips have onboard refresh circuitry. DRAMs are available in 8K, 16K, 64K, 256K, and most recently in 1-meg sizes. K stands for kilobit and refers to the chip's storage capacity. See *kilobit* definition.

Electrically erasable programmable read-only memory (EEPROM): A variety of ROM which can be erased and programmed at the user's factory. The device is similar to the EPROM except it can be erased electrically (without ultraviolet light).

Electrically programmable read-only memory (EPROM): A variety of ROM which can be erased and programmed at the user's factory. The classical EPROM comes with a quartz window in its package so that ultraviolet light can be used to erase its contents. Then each memory location can be programmed to permanently contain desired information. In applications where low volume or time constraints prevent the fabrication of a custom ROM, or where the user may intend to make future modifications to its nonvolatile memory, EPROM devices are used. Sometimes EPROMs are supplied without quartz windows (cheaper). Since ultraviolet light cannot get in to erase these devices, they are programmable only once.

Floating gate: This is the structure in an EPROM device which allows a memory cell to be programmed and later erased. The floating gate can be charged by applying a relatively high voltage to the region surrounding it. Electrical traps in the floating gate store electrons which reach the floating gate. The trapped electrons can be sensed by surrounding structures. When ultraviolet light is directed at the floating gate, the light has sufficient energy to excite the trapped electrons out of the floating gate, and the memory is erased. See *EPROM* definition.

Gallium arsenide: A semiconductor material with properties considered by many to be superior to silicon's. The fastest switching transistors are made with gallium arsenide. Difficulty and expense in device fabrication, as well as constant silicon device improvement, have led to a relatively small market for gallium arsenide products.

Gate oxide: This is a critical part of the MOS transistor which is typically formed by oxidizing the surface of a silicon wafer (to make silicon dioxide) in a high-temperature (1000° C) furnace. The gate itself is typically formed out of a deposited layer of polycrystalline silicon. See definitions for *threshold drift* and *MOS*.

HMOS: An Intel acronym standing for high-performance MOS. HMOS is an NMOS process, with small geometries. See *NMOS* definition.

Kilobit (1K): 2^{10} or 1024 bits. Each DRAM generation has four times as much capacity as its predecessor. Since computers operate in binary code, the actual memory contents are multiples of 2. Thus, the 1K generation has 2^{10} bits, the 4K generation has 2^{12} bits, the 16K generation has 2^{14} bits, the 64K generation has 2^{16} bits, and so on.

Magnetic core: A form of random-access computer memory utilizing ferrite cores to store information. This technology was made obsolete by silicon devices.

Megabit (1 meg): 2^{20} or 1,048,576 bits. See definitions for *kilobit* and *DRAM*.

Metal oxide semiconductor (MOS): Refers to a generic type of transistor (see definition of *transistor*) and to the family of processes used to make it. The switch in an MOS transistor is caused by the action of the metal (or polycrystalline silicon) gate on the "channel." MOS transistors come in two polarities: n-channel (NMOS) or

p-channel (PMOS). To turn on a p-channel device, a negative voltage is put on the gate. The charge on the gate induces an opposite charge in the channel which completes the circuit between the source and the drain. When the voltage is removed, the channel no longer conducts. The n-channel device turns on with a positive voltage applied to the gate. The MOS process typically requires fewer processing steps than the bipolar process. The turn-on speed on MOS devices is controlled by fundamental physics (the mobility of electrons and positive charges in silicon) and the geometry of the device (as devices get smaller, they get faster).

Multiplexing: A generic term used in many areas of electronics. In the case of the 4K and later DRAM generations, multiplexing refers to a scheme adopted to economize on the number of output pins required to address each memory location. Instead of using one pin for each column and each row in the matrix of memory cells, multiplexing allows the 4K memory to be addressed with just 12 pins (it contains 2^{12} bits).

Multitasking: Refers to a microprocessor's ability to manage more than one task simultaneously. Multitasking is not simply a software feature. The ability to employ multitasking is embedded in the chip's architecture.

NMOS: See *MOS*. Several generations of logic were built on NMOS circuitry. A cell of six NMOS transistors replaced Intel's traditional PMOS logic family. NMOS transistors are faster than PMOS devices due to fundamental physical properties.

Plasma etching: A process which is used to define patterns on the silicon wafer during the fabrication process. Until the early 1980s, all etching was done with wet chemicals. Plasma etching improves control and linewidth accuracy. It takes place in a partial vacuum chamber. Gaseous chemicals are introduced into the wafer chamber and ionized using radio frequency power. The ionic species selectively etch different materials used to build the integrated circuit. Plasma chemistry is a new discipline which has been brought to bear on semiconductor processing in order to achieve smaller linewidths and better etching control.

Polycrystalline silicon (poly, polysilicon): A material which can be used as a conductor. In the wafer fabrication process, polycrystalline silicon is deposited on the wafer surface (usually in a low-pressure, high-temperature process) and etched in patterns to form connections between transistors. It is also used to form the gate structure of a transistor (the gate turns the transistor on or off), the floating gate of an EPROM cell (stores the state of the EPROM cell), and one side of the storage capacitor which makes up a DRAM cell. Its main advantage as a material in processing is that it serves as a conductor while also being able to withstand high-temperature processing. While other conductive materials (such as aluminum) cannot withstand the high temperatures required by wafer processing and must be applied only at the end of the process, poly can be applied in the middle of the process and subsequently be covered by other layers.

Polysilicon resistor: By varying the conditions under which polysilicon is deposited on a wafer, lines of polysilicon can be used to form resistor elements. The poly resistor process was difficult for Intel to execute.

Random-access memory (RAM): Formerly called direct-access memory. Family of information storage devices in which specific memory locations can be accessed (to retrieve or store information) in any sequence. This is distinct from sequential-access memory, in which data must be retrieved or stored in a specific order or sequence (example: magnetic tape memory, CCD memory, bubble memory). RAM is usually volatile memory. Thus, a constant power supply is required in order to retain stored information. Several processing technologies have been used to produce the two generic varieties of RAM, DRAM and SRAM.

Read-only memory (ROM): A variety of memory which contains a fixed set of information which cannot be altered, often referred to as nonvolatile memory. Within a typical computer system, ROM contains a sequence of data which has been embedded in the chip at the factory. Thus, ROM chips are custom-made for each application. Only one masking layer in a 10-layer fabrication process needs to be altered to change the information stored in a ROM.

Refresh: Since a dynamic RAM will hold data for only a fraction of a second before it is lost (the charge on the capacitor holds for only a fraction of a second before it leaks away), a useful DRAM must contain circuitry which can continually read and update the contents of each memory location. This circuitry is referred to as refresh circuitry.

Scaling improvements: Refers to the general process of decreasing linewidths in integrated circuits. In the early 80s, Intel's static/logic group focused on taking existing products and shrinking them to improve yield and increase manufacturing capacity. Devices would be shrunk proportionally (nearly), so that chip design would not have to be changed significantly.

Shift registers: A common type of sequential-access memory used in computer systems to manipulate strings of data.

Static random-access memory (SRAM): A RAM memory device which does not require refreshing as long as power is constantly applied. Each memory cell includes either four transistors and two resistors or six transistors. In comparison with DRAMs, fewer memory cells can be packed into the same area. SRAM memory can be made with faster access times than DRAM. The process for SRAM more closely resembles the process for logic devices. As a result, the on-chip memory contained in microprocessors is often SRAM.

Stepper alignment: The latest generation of photolithography processing is carried out on stepper aligners. The photolithography step has two key goals: to align the current mask layer to all previous layers and to transfer the narrowest possible line widths to the wafer. With traditional projection alignment, the pattern for the entire wafer is exposed at the same time. As wafer diameters increase and minimum geometries decrease, the alignment task becomes more difficult. The slightest thermal expansion or warpage will cause the devices on the edge of the wafer to be misaligned even when those in the center are aligned. Stepper aligners expose patterns across the wafer in several steps so that the runoff at the wafer edges can be minimized. At each step, the mask and the wafer are realigned. Stepper aligners are very sophisticated optical and mechanical devices, costing upwards of $1 million per unit.

Threshold drift: Refers to a phenomenon which causes the turn-on voltage of an MOS transistor to change over time. A certain critical voltage must be applied to the gate of an MOS transistor in order to turn it on. If the oxide insulator which separates the gate from the channel is not free of mobile ionic contamination, the threshold, or turn-on, voltage will drift or change over time making the device useless. One source of mobile ionic contamination is common table salt.

Transistor: First invented at Bell Labs in 1948, the transistor is a solid-state device which can be thought of as an electrical switch. It is a three-terminal device: voltage applied to one terminal opens and closes the circuit between the other two terminals. Transistors are the fundamental building block for electronic and logic circuitry. Configurations of transistors can execute logic functions. The first transistors replaced vacuum tubes and were fabricated one at a time by fusing three material layers together in a "sandwich" structure. Bob Noyce (Intel) and Jack Kilby (TI) invented the "planar transistor," which allows fabrication and interconnection of many transistors on one substrate. While many variations exist, two basic types of transistors dominate the current market: bipolar and MOS (or FET) transistors.

Trench etched capacitor: A traditional capacitor is formed on the surface of the silicon wafer (see *capacitor* definition) and occupies a significant portion of a DRAM cell's area. A trench etched capacitor conserves silicon surface area because its orientation is perpendicular to the wafer surface. Vertical trenches are formed using a relatively new technique called reactive ion etching in which the wafer is exposed to a plasma in a strong electric field. Some manufacturers have chosen to adopt the trench structure in order to produce the 1-meg generation of DRAMs. (Note that another method of maintaining storage capacity while reducing area is to reduce the insulator thickness. This has been the traditional method, but has become more difficult in recent generations. Thin-oxide capability is considered a key technological advantage. Current oxide [insulator] thicknesses are about 100 angstroms [one-hundred-millionth of a meter], considered to be near the limit of current manufacturing methods.)

Virtual memory addressing: This microprocessor feature allows the microprocessor to handle many users at the same time without confusing each user's tasks. More specifically, it refers to the microprocessor's ability to use its own protocol to keep track of memory locations regardless of the physical configuration of memory. For example, Intel's 80286 can assign up to one gigabyte of virtual memory addresses to different users. Those virtual memory addresses are then mapped into the physical memory addresses.

Wafer: A slice of silicon which serves as the substrate for integrated circuits. Each wafer contains up to several thousand chips. The first silicon wafers used in production were 2 inches in diameter. Most recently almost all of Intel's fabrication takes place on 6-inch-diameter wafers. In some processing steps such as diffusion, wafers are processed in batches of 25 to 50. Other processing steps such as photolithography take place on individual wafers, one at a time. As processing technology has become more and more complex and wafer size has increased, additional steps have been carried out on individual wafers as opposed to batches.