Lithography and the Future of Moore's Law

Gordon E. Moore
Intel Corporation
2200 Mission College Blvd.
Santa Clara California 95052

The definition of "Moore's Law" has come to refer to almost anything related to the semiconductor industry that when plotted on semi-log paper approximates a straight line. I hesitate to review it's origins and by doing so restrict it's definition.

However, today I will review the history and past performance relative to predictions and show where the advances have come from. I will leave the future performance up to you. Certainly continuing on the same slope doesn't get any easier. It presents a difficult challenge to the industry.

The original paper that postulated the first version of the "law" was an article I wrote for the 35th anniversary issue of Electronics Magazine in 1965. My assignment was to predict what was going to happen in the semiconductor components industry over the next ten years -- to 1975. In 1965 the integrated circuit was only a few years old and in many cases was not very well accepted. There was still a large contingent in the user community who wanted to design their own circuits and who considered the job of the semiconductor industry to be to supply them with transistors and diodes so they could get on with their jobs. I was trying to emphasize the fact that integrated circuits really did have an important role to play.

Let's start with two figures from that original paper. Fig. 1 shows my estimate of the cost of integrated circuits divided by the number of components, a component being a transistor, resistor, diode or capacitor, in an integrated structure at various times. In 1962 the minimum cost per component occurred for circuits containing about ten components. For more complex circuitry costs skyrocketed because yields collapsed. With time, as processing improved, the minimum moved down and to higher complexity. When I wrote this article in 1965 my estimate was that the minimum cost per component was achieved with several tens of components in a circuit, and I predicted that the minimum would continue to go down as we improved out processing capability.

0-8194-1786-6/95/$6.00

TSMC Exhibit 1022, Page 1 of 16

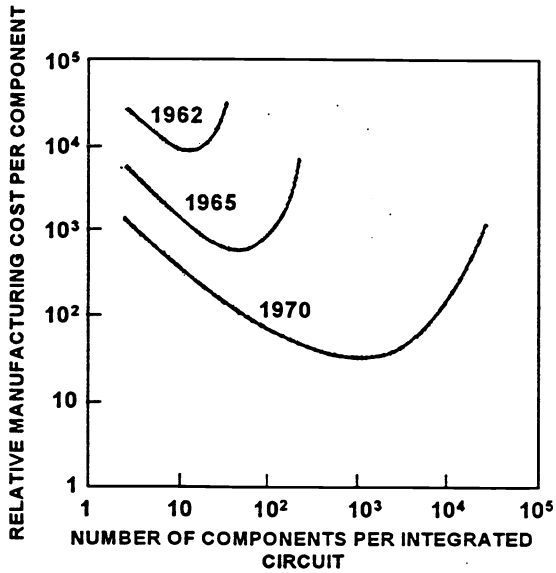

Figure: 1: Relative manufacturing cost per component vs. components in the circuit estimated for various times.

Next I looked at how complex integrated circuits should minimize cost per component and reasoned that the most complex circuit at any time would not be much more complex than this minimum, because of the steepness of the curve beyond the minimum. The available data I chose started with the first planar transistor, which had been introduced in 1959. It was really the first transistor representative of the technology used for practical integrated circuits. It is represented by the first point, two raised to the zero power, or one component.

Adding points for integrated circuits starting with the early "Micrologic" chips introduced by Fairchild, I had points up to the 50-60 component circuit plotted for 1965 as shown in Fig. 2. On a semi-log plot these points fell close to a straight line that doubled the complexity every year up until 1965. To make my prediction, I just extrapolated this line another decade in time and predicted a thousand-fold increase in the number of components in at the most complex circuits available commercially. The cheapest component in 1975 should be one of some 64,000 in a complex integrated circuit. I did not expect much precision in this estimate. I was just trying to get across the idea this was a technology that had a future and that it could be expected to contribute quite a bit in the long run.

TSMC Exhibit 1022, Page 2 of 16

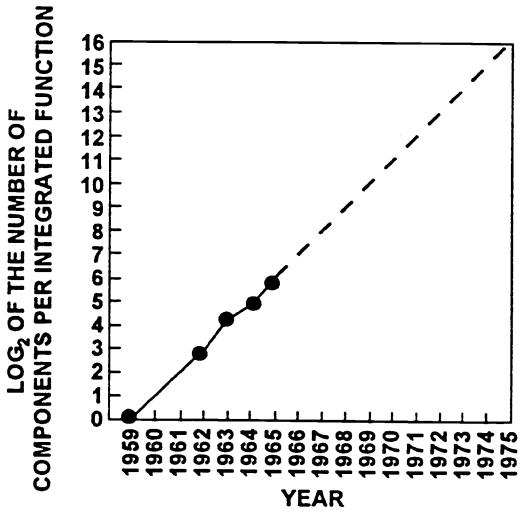

Figure 2. The original "Moore's Law" plot from Electronics April 1965.

Many of you were not in the industry when the devices represented by the first few points in this plot were introduced. I have reproduced photomicrographs of the first planar transistor and the first commercially-available integrated circuit in Figs 3 & 4. I am particularly fond of the transistor, since it is one of the very few products that I designed myself that actually went into production.

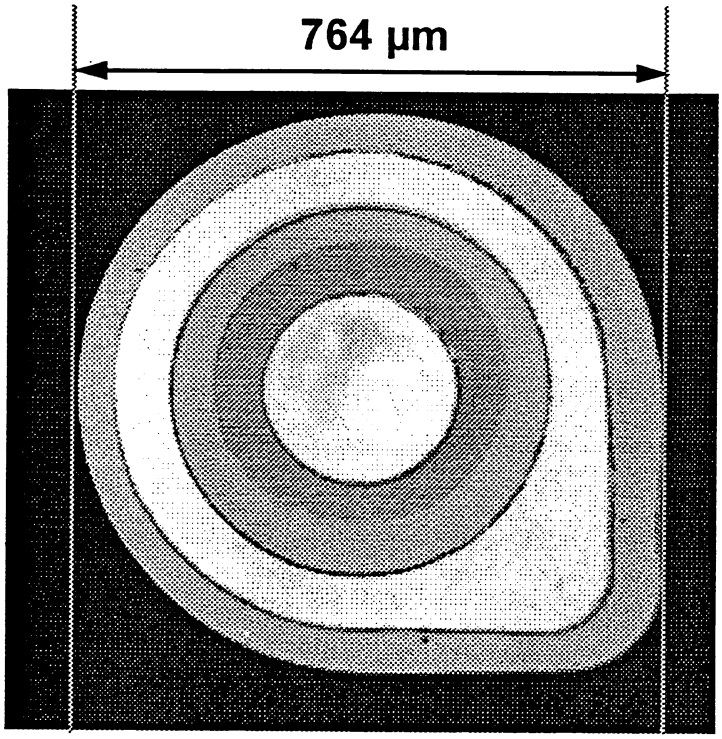

Figure 3. Photomicrograph of the first commercial planar transistor introduced by Fairchild Semiconductor in 1959.

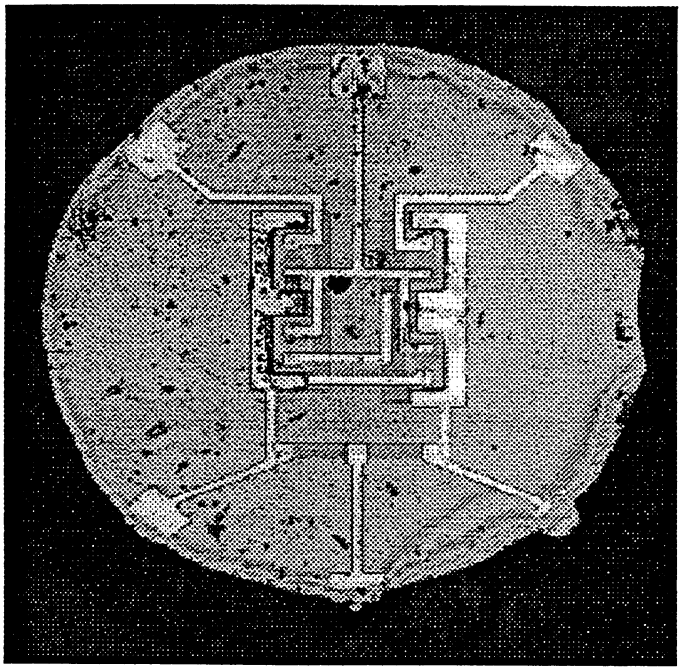

Figure 4. Photomicrograph of one of the first planar integrated circuits produced by Fairchild Semiconductor in the early 1960's.

The unusual diameter of 764 microns was chosen because we were working in English units and that is thirty thousandths of an inch, or 30 mils. The minimum feature size is the three mil metal line making the circular base contact. Metal-to-metal spacing is five mils to allow the 2.5 mil alignment tolerance we needed.

Interestingly enough at the time the idea for the planar transistor was conceived by Jean Hoerni in the early days of Fairchild Semiconductor, it had to sit untried for a couple of years, because we did not have the technology to do four aligned mask layers. In fact, we were developing the technology to do two aligned oxide-masked diffusions plus a mesa etching step for transistors. The original step and repeat camera that Bob Noyce designed using matched 16 mm movie camera lenses had only three lenses, so it could only step a three-mask set. We had to wait until the first mesa transistors were in production before we could go back and figure out how to make a four mask set to actually try the planar idea.

The first integrated circuit on the graph is one of the first planar integrated circuits produced. It included four transistors and six resistors. It has always bothered me that the picture of this important device that got preserved was of the ugly chip shown in Fig. 4. The circuit had six bonding pads around the circumference of a circle for mounting in an 8-leaded version of the old TO-5 outline transistor can. In this case only six of the eight possible connections were required. We did not think we could make eight wire bonds with reasonable yield, so for these first integrated circuits we etched a round die

TSMC Exhibit 1022, Page 4 of 16

that let us utilize blobs of conducting epoxy to make contact to the package pins. For the die in the picture, the etching clearly got away from the etcher.

How good were my predictions? What really happened?

Fig. 5 adds the points for several of the most complex integrated circuits available commercially from 1965 to 1975. It was taken from an update of the industry's progress that I presented at the 1975 IEEE International Electron Devices Meeting. Generally they scatter pretty well along the line that corresponds to doubling every year. For a prediction of a thousand-fold increase in complexity, this fits pretty well. The last point shown, for the most complex device, represents a 16k charge-coupled-device (CCD) memory. I will come back to this in a minute.

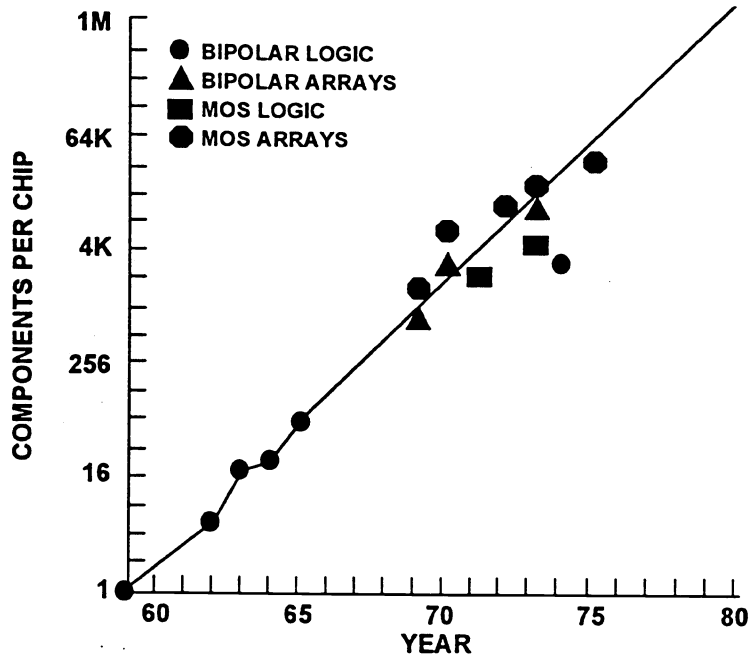

Figure 5. Approximate component count for integrated circuits introduced up to 1975 compared with the prediction of the most complex circuits from the original Electronics paper (Fig. 2.)

This time I tried to resolve the curve into various contributions from the technology to see where the progress was coming from.. First, the dice were getting bigger. As defect densities decreased we could work with larger areas while still maintaining acceptable yields. Many changes contributed to this, not the least of which was moving to optical projection rather than contact printing of the patterns on the wafers. This increase in die area followed a good approximation to exponential growth with time as can be seen from Fig. 6.

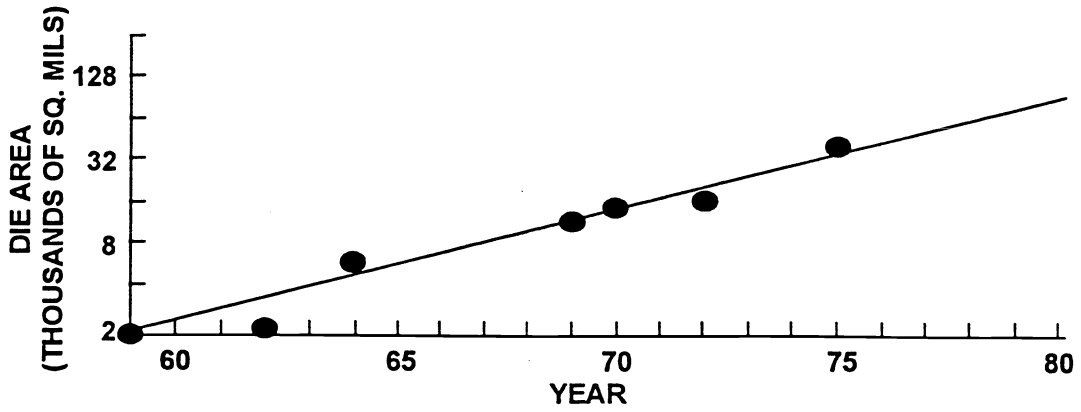

Figure 6. Contribution to increasing complexity resulting from increased die area. Proceedings IEEE IEDM 1975.

Second, not only did we move to larger dice, but we simultaneously evolved to finer and finer dimensions. This increases the density on the chip as the reciprocal of the square of the minimum dimension, or the average of the minimum line width and spacing in cases where they were not equal. Fig. 7 shows that this also approximates an exponential growth of component density if one neglects the first point, which is reasonable, since the planar transistor was not pushed to use the finest features that could be etched. We had enough other problems to deal with on that device.

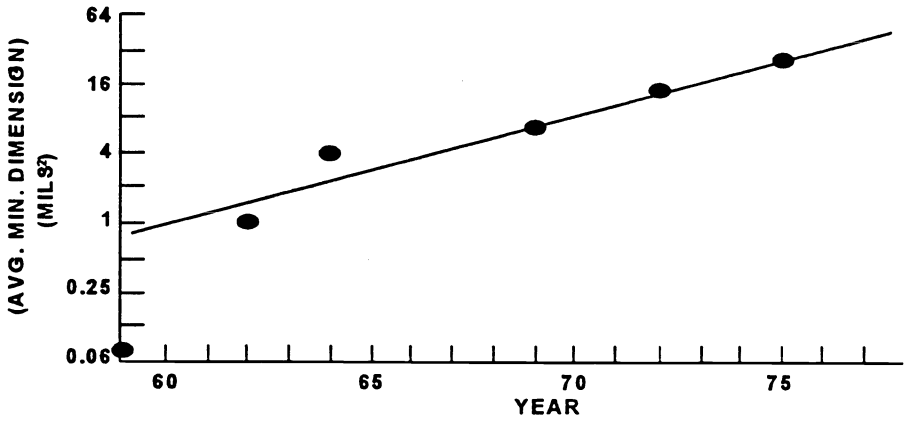

Figure 7. Contribution to increasing complexity from finer structures. Proceedings IEEE IEDM 1975.

TSMC Exhibit 1022, Page 6 of 16

In Fig. 8 the contributions to increased complexity from larger dice and finer dimensions are show along with the increase in complexity achieved. They are shown individually and the product of the two is also shown. Clearly there is another contribution beyond these two. I attributed this additional contribution to "circuit and device cleverness". Several features had been added. Newer approaches for device isolation, for example, had squeezed out much of the unused area. The advent of MOS integrated circuits had allowed even tighter packing of components on the chips.

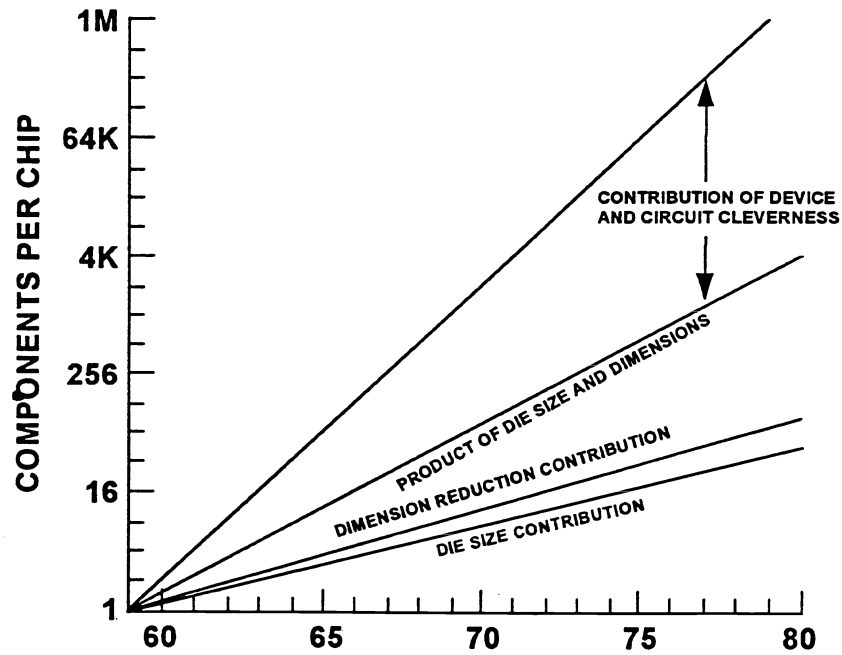

Figure 8. Resolution of the increase in complexity into die size, dimension reduction and "cleverness" factors. Proceedings IEEE IEDM 1975.

Looking at this plot, I said that approximately half the progress had come from die size and finer structures, the remaining half from "cleverness". If one looks closely, however, more than half comes from the first two factors, more like 60 percent. I didn't think that the data was good enough to push this much, however, so I stuck with half.

Remember that my most complex device was a CCD memory. The CCD structure is essentially active device area side by side. There is no room left to squeeze anything out by being clever. Going forward from here we have to depend on the two size factors - bigger dice and finer dimensions.

So I changed my projection looking forward. The complexity curve is going to change slope. Instead of doubling every year, it will be closer to doubling every two years. But as can be seen in Fig. 9., however, I did not predict that the slope would change immediately, but left five years for it to roll over. This delay was because I had too much visibility into what I believed to be the near future.

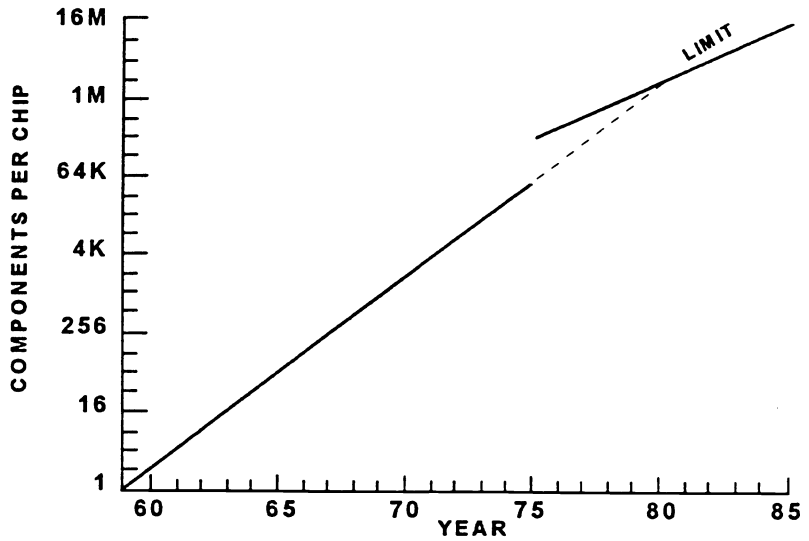

Figure 9. My prediction of the maximum complexity limit in 1975. Proceedings IEEE IEDM 1975.

My problem was the CCD memories. Beyond the 16k that was already on the market, I knew that Intel had a 64k product (>128k components the way I counted CCD's) about ready to announce and we were working on a 256k product. This suggested that I should push out the time at which we slipped from doubling every year, because these CCD's would keep us on the old curve for a while.

The thing that I couldn't know was that the CCD memories would not be introduced. This was just when the soft error problem with DRAMs was discovered, where there would be occasional random losses of bits of information. The cause was traced to alpha particles coming from the packaging materials generating enough hole-electron pairs in the silicon to destroy the charge representing a bit. CCD's are especially sensitive to this phenomenon, in fact that is why they make such good imaging devices. Our CCD's proved to be the best vehicle to study the DRAM problem and to test solutions. As memories, they became just that ... memories. The points that I thought I knew were coming never became commercial products.

TSMC Exhibit 1022, Page 8 of 16

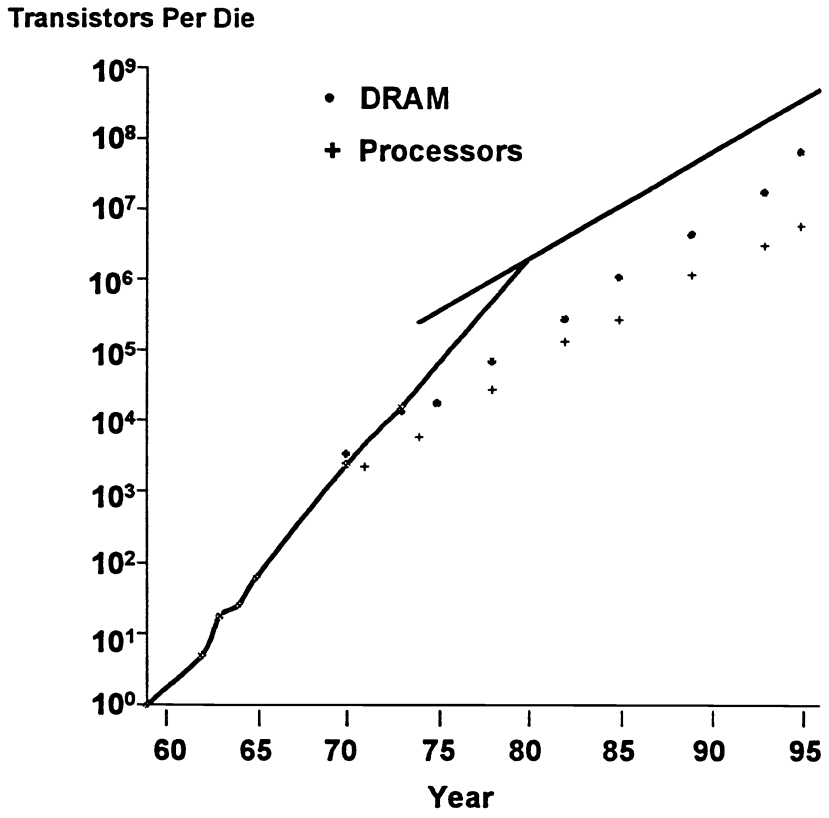

Figure 10. Transistor count for various DRAMs and microprocessors since 1975 compared with the 1975 prediction.

Fig. 10 shows the actual progress in device complexity since 1975 compared with my 1975 prediction. Instead of continuing at the annual doubling rate for the next five years, the slope changes immediately, which is what my calculation said should happen. This figure suggests that the 64meg DRAM will be commercially available this year. If it is delayed to next year, it will fit my plot better. Such a delay would certainly not surprise me.

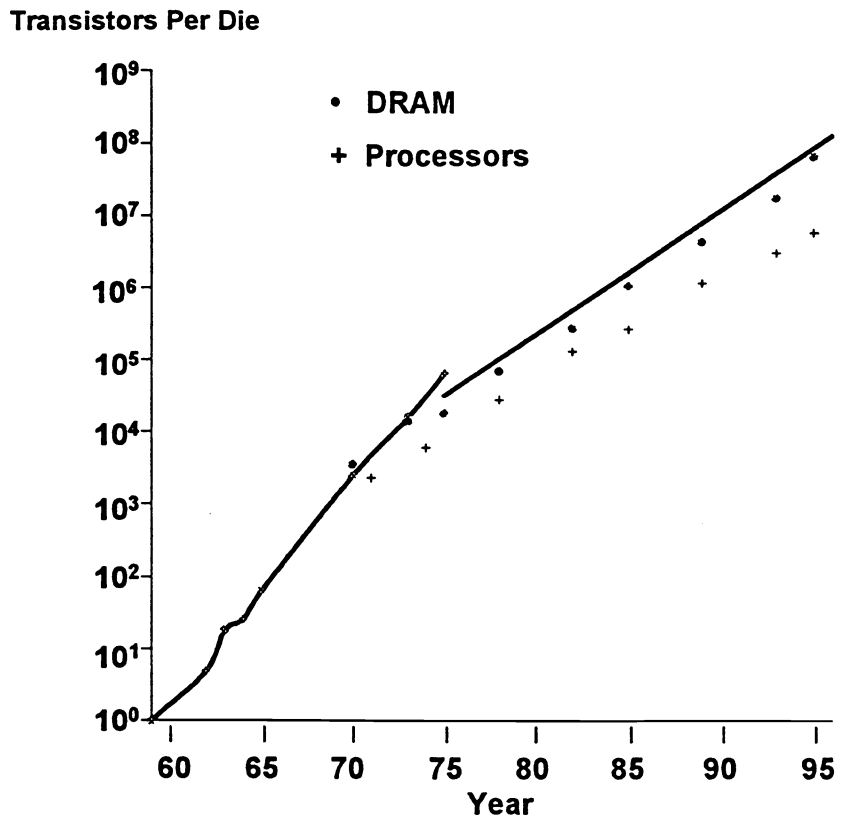

Figure 11. What I calculated and should have said in 1975 vs. what has actually occurred.

Let's look at what I should have said based on my reasoning, rather than what I did say based on too much information.. In Fig 11. the slope changes right away starting with the 32,000 components of the 16k CCD. The new slope is very close to the one predicted by extrapolating the product of die size and density from smaller dimensions, rather than the doubling every two years that I used before If I believed what the data said, I would have drawn that line and then, I guess, I would have had much more reason to be proud of my predictions of what has happened in the last 20 years.

TSMC Exhibit 1022, Page 10 of 16

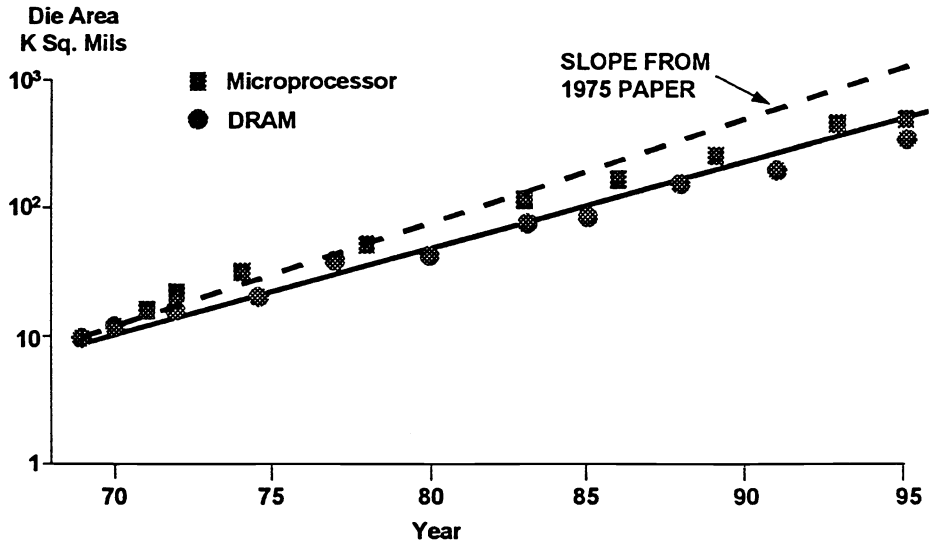

Figure 12. Contribution from increasing die size over the most recent 25 years.

Let's look at how this increase in complexity over the last period can be factored. Amazingly we have stayed very closely on the exponentials that were established during the first fifteen year period. Fig. 12 shows both microprocessor and DRAM die size history. The microprocessor die tend to be a little larger, but the greater number of components is on the DRAMs. The best line through the points has a slightly smaller slope than the one in Fig. 6., but the same within experimental error. We haven't done quite as well here as we were doing during the first ten years.

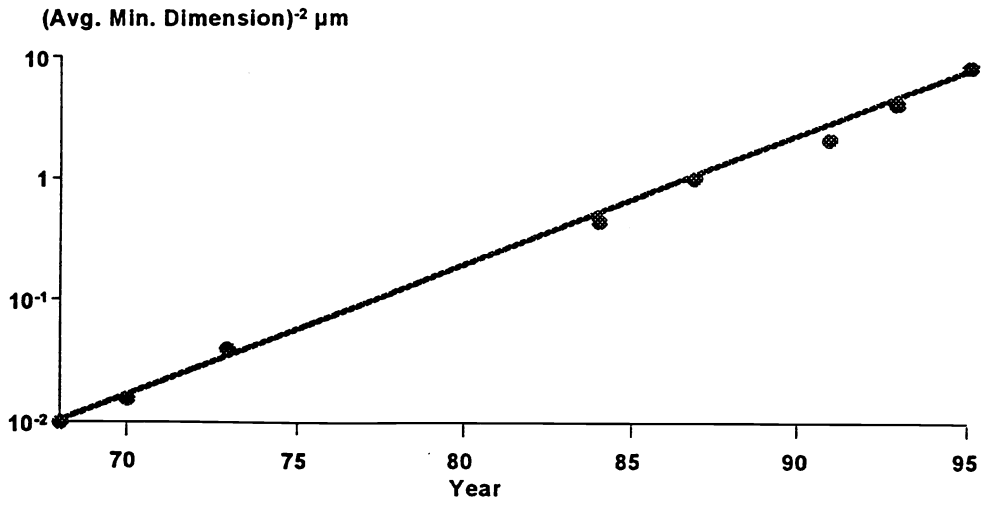

Figure 13. Contribution from increased density from finer line widths over the last 25 years.

On the other hand, looking at Fig 13., the density contribution from decreasing line widths has stayed almost exactly on the same exponential trend as over the first fifteen year period. If anything, the bias is up, but well within the precision of the data. I think that this is truly a spectacular accomplishment for the industry. Staying on an exponential like this for 35 years while the density has increased by several thousand is really something that was hard to predict with any confidence.

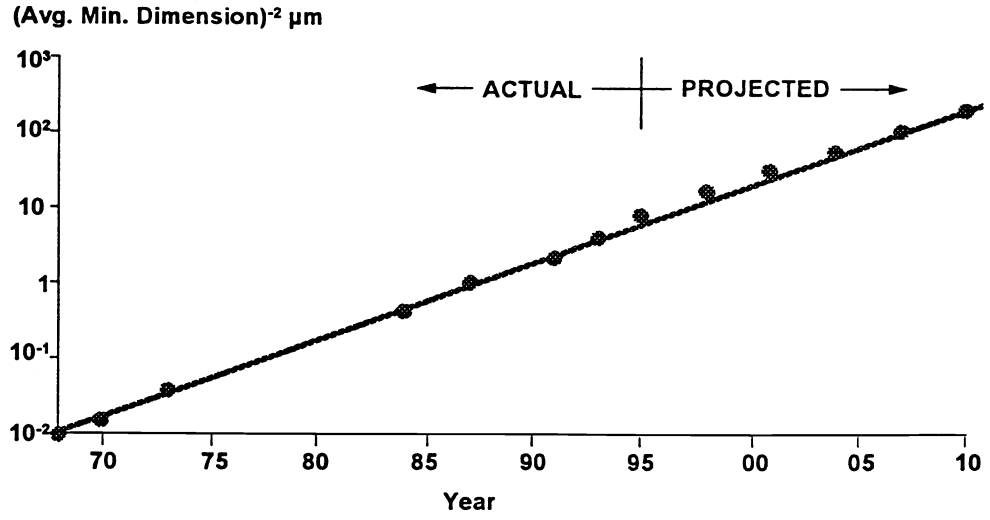

Figure 14. Density contribution historically and extrapolated based upon the SIA technology roadmap through year 2010.

TSMC Exhibit 1022, Page 12 of 16

I have never been able to see beyond the next couple of generations in any detail. Amazingly, though, the generations keep coming one after the other keeping us about on the same slope. The current prediction is that this is not going to stop soon either. The current Semiconductor Industry Assocation (SIA) technology roadmap lays out a path to keep it going well beyond my tenure in the industry (Fig. 14). While I have learned not to predict an insurmountable roadblock, staying on this line clearly gets increasingly difficult. As we go below 0.2 micron, the SIA roadmap says 0.18 micron line widths is the right number, we must use radiation of a wavelength that is absorbed by almost everything. Assuming more or less conventional optics, problems with depth of field, surface planarity and resist technology are formidable, to say nothing of the requirement that overlay accuracy must improve as fast as resolution if we are really to take maximum advantage of the finer lines. These subjects will all be treated in several of the papers at this conference. But what has come to worry me most recently is the increasing cost. This is another exponential, as shown in Fig. 15.

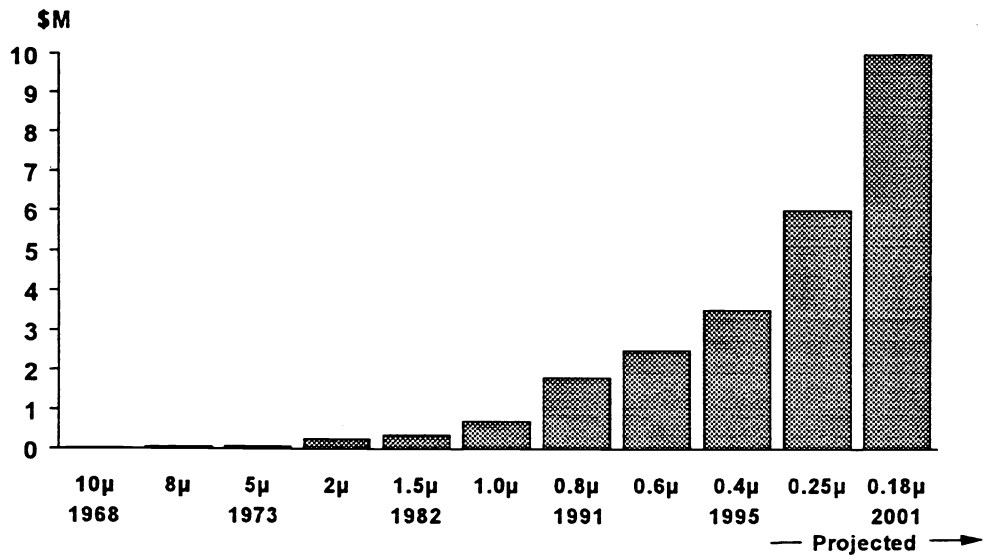

Figure 15. Increasing cost of lithography tools including extrapolation to 0.18 micron minimum dimensions.

When Intel was founded in 1968, we set up our manufacturing facility. A piece of equipment cost about $12,000. You could buy a bank of diffusion furnaces, an evaporator, a lithography exposure machine or whatever for about that amount.

In a couple of years it is going to be $12 million for a production tool. The equipment tends not to process any more wafers per hour than they did in 1968 -- the wafers are bigger, they were two inch then and current eight wafers have sixteen times the area, but that has not gone up nearly as fast as the cost of the equipment has. This is a really difficult trend to stay on. My plot goes only to the 0.18 micron generation, because I have no faith that simple extrapolation beyond that relates to reality. These points are not quotes from equipment vendors, but the best judgment of some of Intel's lithography engineers. Beyond this is really *terra incognita*, taking the term from old maps. I have no idea what will happen beyond 0.18 microns.

In fact, I still have trouble believing we are going to be comfortable at 0.18 microns using conventional optical systems. Beyond this level, I do not see any way that conventional optics carries us any further. Of course, some of us said this about the one micron level. This time, however, I think there are fundamental materials issues that will force a different direction. The people at this conference are going to have to come up with something new to keep us on the long term trend.

If one takes the increasing cost of production tools combined with the increasing number of layers in advanced technologies, the cost for a reasonably balanced production facility (about 5,000 wafers per week) grows as shown in Fig. 16. The 0.25 micron plants

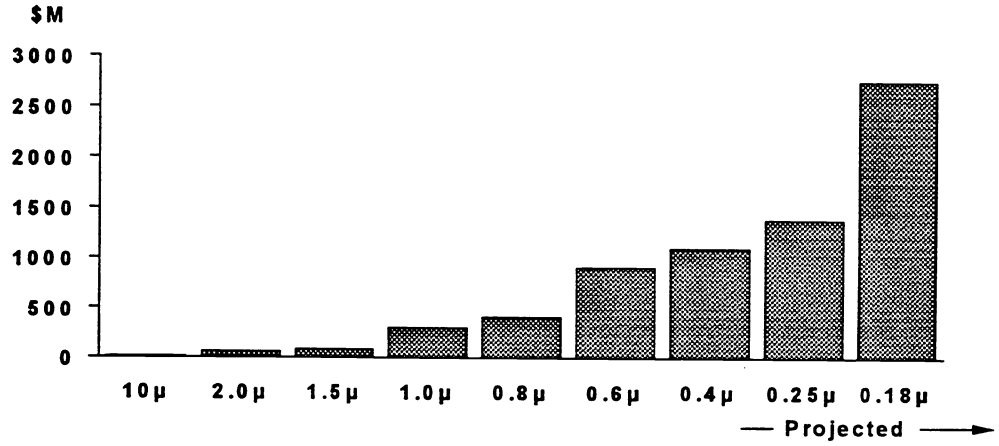

Figure 16. Estimated cost of a wafer processing plant and equipment for 5,000 wafer starts per week for various generations of technology.

are already being constructed as the process is being developed. We have passed the days of mere billion dollar plants. Current facilities under construction will exceed two billion and the three billion dollar plant starts construction no later than 1998. The rising cost of the newer technologies is of great concern. Capital costs are rising far faster than revenue in the industry. We can no longer make up for the increasing cost by improving yields and equipment utilization. Like the "cleverness" term in device complexity disappeared when there was no more room to be clever, there is little room left in manufacturing efficiency.

TSMC Exhibit 1022, Page 14 of 16

Increasing the growth rate of the industry looks increasingly unlikely. We are becoming a large player in the world economy. Fig. 17 shows the semiconductor industry compared with the sum of the gross domestic products of the countries of the world, the Gross World Product would be an appropriate name for it. Obviously this extrapolation has some problems associated with it.

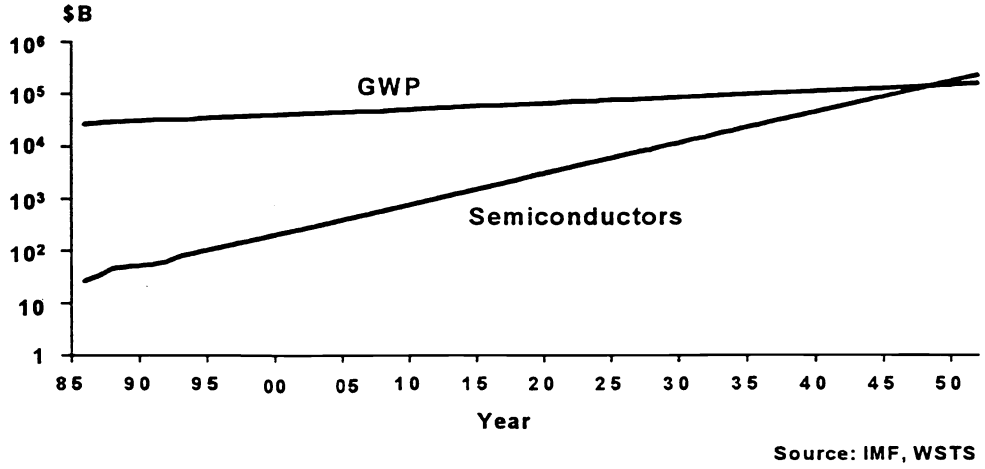

Figure 17. World output of goods and services compared with historic semiconductor industry extrapolated into the future.

As you can see, in 1986 the semiconductor industry represented about 0.1 percent of the GWP. Only ten years from now, by about 2005, if we stay on the same growth trend, we will be 1%; and by about 2025, 10%. We will be everything by the middle of the century. Clearly industry growth has to roll off.

I do not know how much of the GWP we can be, but much over one percent would certainly surprise me. I think that the information industry is clearly going to be the biggest industry in the world over this time period, but the large industries of the past, such as automobiles, did not approach anything like a percent of the GWP. Our industry growth has to moderate relatively soon. We have an inherent conflict here. Costs are rising exponentially and revenues cannot grow at a commensurate rate for long. I think that this is at least as big a problem as the technological challenge of getting to tenth micron.

I am increasingly of the opinion that the rate of technological progress is going to be controlled from financial realities. We just will not be able to go as fast as we would like because we cannot afford it, in spite of your best technical contributions. When you are looking at new technology, please look at how to make that technology affordable as well as functional.

TSMC Exhibit 1022, Page 15 of 16

Our industry has come a phenomenal distance in what historically is a very short time. I think our progress to a considerable extent is the result of two things: a fantastically elastic market with new applications that can consume huge amount of electronics, and a technology that exploits what I have often described as an exception to Murphy's Law.

By making things smaller, everything gets better simultaneously. There is little need for tradeoffs. The speed of our products goes up, the power consumption goes down, system reliability, as we put more of the system on a chip, improves by leaps and bounds, but especially the cost of doing thing electronically drops as a result of the technology. Today one can buy a four megabit DRAM with well over four million transistors on the chip for less than the planar transistor pictured in Fig. 3 sold for in 1960, even neglecting the change in the value of the dollar. We have made of the order of a ten million fold decrease in the cost of a transistor and thrown in all the interconnections free, using the DRAM as an example. It is hard to find an industry where the cost of their basic product has dropped ten million fold even over much longer time periods.

The only one I can find that is remotely comparable is the printing industry. Carving a character into a stone tablet with a chisel probably cost quite a bit, maybe the equivalent of a few dollars today based on the time it probably took. Today people printing newspapers sell them for a price that makes the individual characters about as expensive as are individual transistors in a DRAM.. Surprisingly, they sell about as many characters as we sell transistors, as near as I can estimate. Trying to estimate the number of characters printed is far more challenging than estimating the number of transistors produced. Taking into account newspapers, books, the Xerox copies that clutter up your desks, all such printed matter I estimate that it is no more than an order of magnitude greater than the number of transistors being produced.

Printing in advancing it's technology over the centuries has had a revolutionary impact on society. We can now archive our knowledge and learn from the collected wisdom of the past increasing the rate of progress of mankind.

I think information technology will create its own revolution in society over a much shorter time scale, primarily because of the semiconductor technology you are driving. Semiconductor technology has made its great strides as a result of ever increasing complexity of the products produced exploiting higher and higher density to a considerable extent the result of progress in lithography.

As you leave this meeting I want to encourage each of you to think smaller. The barriers to staying on our exponential are really formidable, but I continue to be amazed that we can either design or build the products we are producing today. I expect you to continue to amaze me for several years to come.

TSMC Exhibit 1022, Page 16 of 16