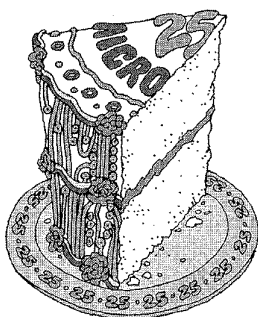# THE FUTURE OF MICROPROCESSORS

Albert Yu

Intel Corporation

Intel's head of microprocessor products looks 10 years ahead to 2006.

In my role as head of Intel's microprocessor products, I am often asked to paint a picture of the microprocessor of the future. Even if our newest processor has just hit the streets and has not even come close to full use, people naturally crave information about where they're going rather than where they've been.

My colleagues and I have been trying for about 10 years now to identify trends about the microprocessor of the future. While these are based on a wide variety of unknown factors inherent in developing new technology, for the most part, we have been close to the mark. However, before making statements about microprocessor trends 10 years out—Micro 2006—it might be useful to revisit our past statements[1,2] about the microprocessor of today and the microprocessor of 2000. Then we can see where we have been right and where wrong. This retrospective will reveal important trends that promise to give some insight into the microprocessor of the next decade.

Performance, capital costs

Over the last 10 years, evolving microprocessor performance increased at a higher than envisioned rate; unfortunately, so did manufacturing capital costs. Table 1 lists our 1989 predictions for today's microprocessor performance at speeds of 100 MIPS (millions of instructions per second), which is equivalent to an ISPEC95 rating of 2.5 and clock rates of 150 MHz. Surprisingly, today's performance dramatically exceeds this. The Intel Pentium Pro processor runs at 400 MIPS, with an ISPEC95 rating of about 10 and a 200-MHz clock rate. This great performance boost has stimulated a huge range of applications for business, home, and entertainment, from mobile computers to servers. As a result, the PC market segment is a lot larger today than we anticipated years ago.

The bad news is that producing advanced microprocessors involves much higher capital cost than anyone ever expected. At Intel, we've augmented Moore's law (the number of transistors on a processor doubles approximately every 18 months) with Moore's law 2. Law 2 says that as the sophistication of chips increases, the cost of fabrication rises exponentially (see Figures 1 and 2). In 1986, we manufactured our 386 containing 250,000 transistors in fabs costing $200 million. Today, the Pentium Pro processor contains 6 million transistors but requires a $2 billion facility to produce.

Looking ahead, the important technological fact that emerges is that Moore's law continues to reign, with the number of transistors per chip increasing exponentially. Today's performance trend can continue, thanks to microarchitecture and design innovations beyond raw transistor count. The personal computer market, by far the biggest market for microprocessors, continues to grow at a healthy rate. It can provide the volume markets needed to absorb the huge manufacturing capital costs. To be sure, we have a number of key technology barriers to overcome as device geometry migrates well below the submicron range. However, all indications are that the microprocessor of 2006—and beyond—will be well worth the wait.

Micro 2000 revisited

As Table 1 shows, we anticipated in 1989 that in 2000 a processor would carry 50 million transistors in a 1.2-in. (square) die. The industry is mostly on track to deliver a 40-million-transistor chip on a 1.1-in. die in 2000. This 20 percent offset is not a technology limitation but an economic one, necessitated by creating a reasonable die cost (see Figure 1).

Silicon technology. Our visions about

46 IEEE Micro

0272-1732/96/$5.00 © 1996 IEEE

silicon process line width were right on the money, as Intel is currently in production with 0.35-micron technology for the Pentium and Pentium Pro. I believe that line width will continue to drop to 0.2 micron in 2000 and to 0.1 micron in 2006 (see Figure 3). Also, the dielectric thickness and the voltage supply will have decreased correspondingly. This incredible shrinkage will continue unabated for the forseeable future. The number of metal interconnects has increased from two to five over the last 10 years and will increase further as we need more interconnects to hook up all the devices.[3] In fact, this is one of the biggest performance-limiting factors we contend with (see later discussion).

In addition, the problem of interconnects from the chip to the package and eventually to the system board is another major limiting factor for performance. Actually, we want to build single chips to avoid performance loss when sending signals off chip. We added cache and floating-point units on the 486 processor

Table 1. Visualizing trends for the microprocessor of the future.

Characteristic	1989 predictions for 1996	1996 actuals	1989 predictions for 2000	1996 predictions for 2000	1996 predictions for 2006
Transistors (millions)	8	6	50	40	350
Die size* (inches)	0.800	0.700	1.2	1.1	1.4
Line width (microns)	0.35	0.35	0.2	0.2	0.1
Performance:					
MIPS	100	400	700	2,400	20,000
ISPEC95	2.5	10	17.5	60	500
Clock speed (MHz)	150	200	250	900	4,000

*Length of single side of square die.

Figure 1. Chart showing Moore's law.

mostly for that reason. For the Pentium Pro processor, we placed the second-level cache and the processor in the same package to achieve the bandwidth needed between the two. The future trend will be to incorporate more performance and bandwidth-sensitive elements on chip and to continuously improve the package interconnect performance. Several companies are investigating MCM (multichip module) technology to eliminate chip packaging altogether, and I believe this will be an important trend for future high-performance processors.

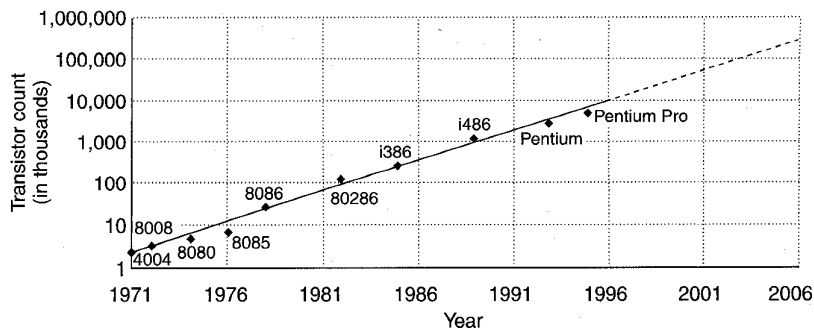**Performance.** It is amazing that the actual performance of microprocessors exceeds our 1989 vision by quite a lot. There are several reasons for this. Although the silicon process advances were pretty much on target, we have achieved higher frequency out of these advances with novel microarchitecture and circuit techniques. In addition, the number of instructions per clock has increased faster, and we have exploited superscalar architectures and greater degrees of parallelism. There have also been significant innovations in compiler technology that boost performance even higher. I see these trends continuing.[4,5]
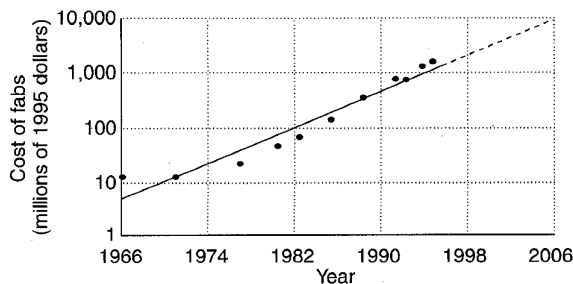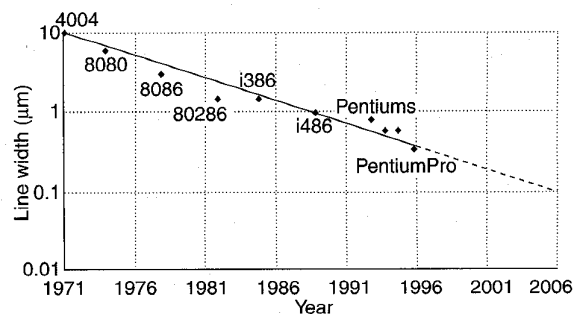

Figure 2. Chart showing Moore's law 2.

Figure 3. Chart showing line width versus time.

Scaling of MOS technology
Carver A. Mead, California Institute of Technology

The MOS transistor is the workhorse of modern microelectronics. Reducing the feature size of CMOS fabrication processes has been the primary method by which ever-increasing computation could proceed at ever-decreasing cost and power consumption. How does this scaling affect device performance? Are there fundamental physical limits to how small the MOS device, as we know it today, can be scaled?

Transistor current is the flow of mobile channel charge induced by an equal charge on the gate. For a logic circuit, the supply voltage induces the channel charge, creates an electric field in the channel, and is the difference between output logic levels.

In long-channel devices, the charge velocity is proportional to the electric field in the channel. The channel current is the product of the channel charge and velocity. Therefore, the device current has a quadratic dependence on the supply voltage. This current must charge the load capacitance to approximately one-half of the supply voltage to achieve a logic transition. Thus, circuit speed is linear in the supply voltage—a dependence that kept power-supply voltages artificially high until a few years ago.

For device dimensions below 1 micron, the old scaling dependence no longer holds—charge velocity becomes independent of electric field. Decreasing the supply voltage no longer decreases the channel current. The same factor decreases both the output current and output voltage. In this regime, the only effect of decreased supply voltage is a decrease in the switching energy, with virtually no decrease in performance.

It is imperative to reduce the supply voltage for reasons other than reducing power consumption. To induce sufficient charge in the channel with a lower operating voltage, we must further thin the gate oxide. The sum of the source and drain depletion-layer thicknesses must be less than the channel length. It is inevitable that, as "minification" continues, these dimensions will become sufficiently small that electron tunneling through them will become comparable to other device currents. These parasitic currents are exponential functions of the supply voltage.

I have presented these considerations earlier. The most remarkable conclusion of my work is that transistors with 0.03-micron channel lengths will operate on a 0.14-volt power supply about three times faster than do today's best devices. Only below this scale do parasitic currents overwhelm the energy consumed in the performance of real computation.

The enormous effect of device scaling on computational capability becomes apparent only when viewed from the system level. We'll see systems integrated to upward of 10^9 devices per square centimeter. Interconnects—both within a single chip and across chip boundaries—determine the dominant signal latency. Even today, it has become more economical to break each chip into several processors that can operate in parallel than to build larger "dinosaur" processors.

Massive parallelism is possible in present-day technology; it will become mandatory if we are to realize even a fraction of the potential of more highly evolved technology. Each processor can operate with its own local synchronous timing, with self-timed signaling between processors.

We have never been able to see more than about two technology generations ahead. In spite of our myopia, *the technology will continue to evolve.* It will evolve because that evolution is possible; because we gain so much at the system level by that evolution; and because the same energy and will on the part of bright, energetic, devoted people that have overcome enormous obstacles in the past will overcome those that lie ahead.

Carver A. Mead *is the Gordon and Betty Moore professor of engineering and applied science at the California Institute of Technology in Pasadena, California. He works on VLSI design, neuromorphic systems, and the physics of computation.*

Reference
1. C. Mead, "Scaling of MOS Technology to Submicrometer Feature Sizes," J. VLSI Signal Processing, Vol. 8, 1994, pp. 9-25.

We will see clock speeds of about 900 MHz with a 60 ISPEC95 rating in 2000. Such tremendous clock rates place great demands on the resistance and capacitance of the chip's metal interconnects for power and clock distribution. These multimillion-transistor devices also face new hurdles in packaging and power management.

Architecture. In the late 1980s, there was much debate about which microprocessor architecture held the key to fastest performance. RISC (reduced instruction set computing) advocates boasted faster speeds, cheaper manufacturing costs, and easiest implementation. CISC (complex instruction set computing) defenders argued that their technology provided software compatibility, compact code size, and future RISC-matching performance.

Today, the architecture debate has pretty much become a nonissue. Both the debate and the competition have been good for the industry, as both sides learned a great deal from the other, which stimulated faster innovation. There is really no perceptible difference between the two in either performance or cost. Pure RISC chips like the IBM ROMP, Intel 80860, and early Sun Sparc, as well as pure CISC chips like the DEC VAX, Intel 80286, and Motorola 6800, are gone. Smart chip architects and designers have incorporated the best ideas from both camps into today's designs, obliterating the differ-

ences between architecture-specific implementations. What counts most in designing the highest performance, lowest cost chip today is the quality of implementation.

Seven years ago in *IEEE Spectrum*,[1] our vision was that the microprocessor of 2000 would have multiple general-purpose CPUs working in parallel. What has instead happened is not separate CPUs on the same chip but a greater degree of parallelism within a single chip. The Pentium processor employs a superscalar architecture with two integer pipes, and the Pentium Pro processor design expanded that to three. Other processors such as the HP PA and IBM PowerPC have used similar superscalar architectures. I see the trend to exploit more parallelism continuing well into the future.

Human interface. The number of transistors devoted to the human interface is increasing too. Human interface functions are those that contribute to making a PC or other device more attractive and easier to use—three-dimensional graphics, full-motion video, voice generation and recognition, and image recognition. Even though we have no way of knowing precisely how future microprocessors will be used, I firmly believe that graphics, sound, and 3D images will play a huge role. We live in a 3D color world, and we naturally want our computers to mirror that. Once the computing power is available to create these kinds of features, application developers will have a huge opportunity to push computing into new realms. Therefore, we'll see a higher percentage of the microprocessor chip allocated to these purposes.

In 1989, we set aside 4 to 8 million transistors—roughly 10 percent of our estimate for 2000—for human interface and graphics functions. Our new MMX technology for the Pentium processor and Sun UltraSparc's VIS (visual instruction set) are examples of general-purpose instructions for accelerating graphics, multimedia, and communication applications.

Bandwidth. What becomes very apparent in moving into the future with complex chips is that microprocessor design is becoming system design. The microprocessor designer must consider everything that touches the chip, which includes the system bus and I/O, among others. As raw processor speed increases, system bandwidth becomes more critical in preventing bottlenecks. We will need very high bandwidth between the CPU and memory and between other system components to deliver the kinds of real speed gains of which the silicon is capable. Toward that end, microprocessor buses continue to increase in throughput. PCI is one of the major standards that allows PCs to increase the I/O bandwidth significantly.

Today, Intel is working with the PC community to spearhead the development of the accelerated graphics port (AGP). This vehicle increases bandwidth between the graphics accelerator and the rest of the system. The AGP will be critical for the full fruition of applications involving 3D and other high-resolution graphics. As communications become even more important for PCs and Internet applications expand, we will need more communications bandwidth.

Design. We saw that our dependence on advanced computer-aided design tools would soar, and it has. Today, we're simulating an entire chip, rather than just portions of it, from behavior to the register-transfer level. CAD tools assist in the entry of various circuit-logic data, verify the global chip timing, and extract the actual layout statistics and verify them against the original simulation assumptions. One of the rapidly developing areas is synthesis, first in logic synthesis but progressing to data path synthesis. These capabilities have improved design productivity enormously.

Future advances will improve the layout density (to reduce product cost) and raise performance (to enable new applications). This is particularly challenging as interconnects are becoming greater performance limiters than are transistors. In addition to electrical simulation, thermal and package simulation will be the norm by 2000. Beyond the chips, the trend is to expand simulation to encompass the whole system, including processor, chip sets, graphics controller, I/O, and memory.

Though the dependency on and rapid innovations in CAD have been pretty much on target, the design complexity and design team size have grown greater than expected. Two engineers developed the first microprocessor in nine months. Modern microprocessor design requires hundreds of people working together as a team.

Though design productivity has improved enormously, it is just barely keeping up with the increased complexity and performance. Looking forward, I see that one of our most challenging areas is how to achieve quantum leaps in design productivity. An obvious help would be for CAD tools to be truly standards-based and fully interoperable. This is not the case today, causing the industry to waste valuable resources struggling with conflicting and proprietary interfaces.

Testing. Testing complex microprocessors has become a huge issue. Though the capital associated with testing microprocessors is still smaller than that associated with wafer testing, it has been escalating beyond our anticipations. Why? First, the tester is more expensive due to increased frequencies and the large number of pins (the Pentium Pro processor runs at 200 MHz and has 387 pins). Second, testers that previously cost $50,000 cost well over $5 million today. Lastly, because of chip complexity and quality requirements of less than 500 DPM (defects per million), test time continues to increase. As a result, the total factory space and capital costs devoted to test have skyrocketed.

In 1989 we envisioned that a larger share of transistors in 2000 would be devoted to self-test—approximately 3 million transistors (6%) out of the total 50 million. A great deal of innovation has happened in this area. Today, roughly 5 percent of the Pentium Pro processor's total transistor count supports built-in self-test. Therefore, our prediction for 2000 stands: About 6 percent or so will be devoted to testing; this number may increase in 2006.

Compatibility. We posited in 1989 that binary compatibility was absolutely critical for investment protection and continuity. There are vast software bases in use today that each year become more valuable assets to businesses. Companies do not want to abandon these, even in favor of faster computers. Thus, even with fairly radical architectural departures such as massively parallel processing, we must maintain compatibility between future microprocessors and today's microprocessors. Only a twofold or greater improvement in system performance makes a switch to incompatible hardware worthwhile. This is more true today than ever
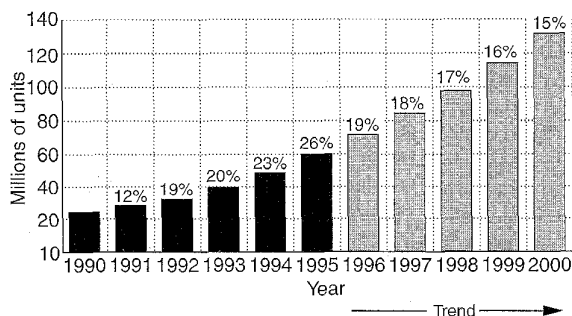
Figure 4. PC shipment trend. *(Source: Dataquest, Apr. 1996)*

and will continue to be one of the most important business and user requirements for future microprocessors. Of course, software is becoming more portable, but no one will devote the resources to recompile and maintain another binary version without major added benefits.

At the same time, the task to ensure compatibility has grown enormously. The number of different operating systems, applications, and system configurations has skyrocketed beyond earlier estimations. Of course, this job of compatibility validation is much harder after the silicon stage than before it, but accomplishing the technical problems with sufficient speed on software models or hardware emulators is an enormous task.

Market segment size. When we had the Pentium processor on the drawing board, we were anticipating sales of only about three million units in 1995. According to IDC reports, Pentium processor shipments in 1995 were close to 60 million. This twentyfold jump has been great for the whole industry. For example, Figure 4 shows Dataquest's estimation of PC shipments through 2000 predicting steady growth of 15 to 19 percent. Lucky for all of us, this market segment growth will allow more R&D dollars and capital investments to drive the microprocessor evolution at the exponential pace of Moore's law.

What about 2006?

Once we understand where we are versus our earlier vision, it is easier for us to look 10 years ahead to 2006.

Transistor and die size. Table 1 and Moore's law show that the number of transistors could jump to about 350 million in 10 years. Remember that plenty of previous-generation processors will continue to ship in huge volumes.

Die size will push toward 1.4 inches to accommodate the tremendous number of transistors and interconnects. Line width will have shrunk to a mere 0.1 micron, stretching today's optical systems to the physical limits. We may well have to look for other alternatives. Silicon technology will continue to advance at a rapid rate, as predicted by Moore's law, and voltage will continue to shrink to well below 1 V.

Performance and architecture. By 2006, performance will have jumped to an incredible 4 GHz or a 500 ISPEC95 rating.[6] All indications are that more opportunities exist for

innovation in performance than ever before. The two trends driving increased performance will continue to be more parallelism and higher frequencies. To exploit more parallelism, we will increasingly focus on compiler and library optimization. To push to higher frequencies, we will need advances in microarchitecture, circuit design, accurate simulation, and interconnects.

I see a great many good ideas that can be implemented for years to come. The performance drive is clearly not bound to the microprocessor but derives from the whole system, as one must build balanced systems to deliver power to users. Interestingly, earlier microprocessors borrowed lots of good architectural ideas from mainframes. From here on out, we are going way beyond the performance any mainframe has ever provided. Therefore, it is also important that the industry devotes more resources to long-term research and forges stronger cooperation with universities.

Barriers. Before we can realize a microprocessor of this complexity, we'll need to meet and resolve several technological and logistical barriers. One of the most basic is grappling with design complexity and the burgeoning size of the design team. Larger design teams are harder to coordinate and ensure communication within. Designing for correctness from the beginning remains a necessity, but becomes far more difficult as designs become exponentially more complex.

Compatibility validation becomes unbelievably difficult in designs as complex as the one we are contemplating for 2006. The task of exhaustively testing all possible computational and compatibility combinations is huge. We need a breakthrough in our validation technology before we can enter the 350-million-transistor realm.

Another area crying out for breakthrough thinking is power. Faster microprocessors obviously need more power, but we also need a way to dissipate the power from the chip through the package and the system. To lower on-chip power, we need breakthroughs to drive voltage requirements way below 1 V. We need innovations in low-power microarchitectures, design, and software to contain the power rise. For mobile applications, the whole electronics complex needs to stay below 20 W. Power poses big challenges not only to microprocessors but to other components in the system such as graphics controllers and disk drives.

As mentioned earlier, interconnects are the major performance limiter and will remain so until scientists discover lower resistance, lower capacitance materials. Today's Pentium Pro processor has five metal layers; future generations will need more. Metallization technologies historically take years to develop, so we urgently need research in this area to create the microprocessor of 2006.

Market segment. We have historically erred on the side of underestimating microprocessor demand. Although I cannot estimate the exact volumes, I do foresee strong continued growth for the PC and microprocessor market segments into the next decade. Although the PC market segment in the United States is maturing, it is just beginning in emerging markets, notably Southeast Asia, South America, and Eastern Europe.

In addition to openings of new geographical markets, new

Mediaprocessors

John Moussouris, MicroUnity Systems Engineering, Inc.

Microprocessors have evolved over the last quarter century as self-contained devices for calculating and controlling things. Growth in electronics is now shifting toward interconnected devices whose primary function is to communicate. The goal of delivering the content and services of the entire global network with an ease and affordability more like TV, radio, and telephone than personal computers will impact processor evolution enough to merit a distinct category: the mediaprocessor.

How do communication algorithms differ from classical embedded and desktop applications? Classical applications typically perform arithmetic, Boolean, and shift operations on a few different data sizes (for example, 32-bit integers; 64-bit floating point). Communication processes, on the other hand, operate on a much wider range of data widths and mathematical domains (such as Galois fields used to compute generalized parity or "syndromes").

Encryption and error-correction codes require bit-level and Galois processing. Video, RF, and modem processing need 2 to 12 bits to represent their samples; audio commonly uses 8 to 24 bits; and packet protocols need thousands of bits. A single sample may require hundreds or thousands of operations in the course of filtering, compression, encryption, modulation, transmission, equalization, demodulation, and error correction. Such high broadband rates strain both computational throughput and bandwidth of the memory system. On the other hand, the total memory required is often small—typically dominated by a few megabytes of frame storage.

Legacy microprocessor architectures have had to respond to these needs. Most have defined multimedia extensions that improve support for audio, video, and graphics. They also are incorporating interfaces for faster memory and real-time I/O. Backward compatibility with legacy code and interfaces, however, adds complexity and cost to these designs.

New mediaprocessor technologies aim to reduce this cost. One area of innovation is execution units that systematically and efficiently implement subinstruction level parallelism such as SIGD (single instruction on groups of data) over all multiprecision data types. Another area is programming models that eliminate redundant register files, condition codes, and mode bits. These models simplify code generation and streamline interlocks, bypass, and exceptions in pipelined and instruction level parallel machines (VLIW, superscalar, and decoupled access execution designs). A third is efficient protection and synchronization mechanisms for the sharing of memory and data path resources among many user level and secure-kernel threads of execution, enabling thread level parallelism. A fourth is packet-oriented interfaces compatible with multiple streams of broadband traffic across few packages at low pin count.

Most of these innovations harness parallelism inherent in communications. The winning approach is modest use of each of these bandwidth-enhancing mechanisms in a mathematically pure and concise architecture. For example, a mediaprocessor with 128-bit, SIGD, 4-operand instructions, four-way issue, and five threads would achieve about 10,000 bits of operand throughput per cycle. This is compatible with very low voltage or the small driver operation needed to save power and silicon area. Efficient uses of these degrees of parallelism, moreover, are within reach of current vectorization and instruction- and thread-scheduling software technology.

Ultimately, the dominant cost in broadband evolution will be the development and maintenance of an enormous body of software. Current microprocessor hardware and software is stretching to accomplish the audio, video, graphics, and GUI processing needed at the presentation and application layers of the communications protocol stack.

Far greater challenges remain at the lower transport to physical layers, where algorithms are evolving to enable broadband and wireless links in the network. The high standards of code robustness needed for these lower layers are inspiring new CASE methodologies, such as symbolic verification. The greatest economy in media processing will derive from amassing rich software development tools around a general and unified programming model that supports the entire communications protocol suite.

John Moussouris is chair and CEO of MicroUnity Systems Engineering, Inc., in Sunnyvale, California, a developer of mediaprocessor architecture and software. Previously, he cofounded and served as vice president of VLSI Development at Mips, where he led the architecture and initial implementation of the Mips microprocessor line. He began his career as a research staff member in VLSI and RISC design at the IBM T.J. Watson Research Center.

functional markets will continue to unfold. Although it is the futurist's job to imagine how computing power will be used in the next century, history shows that incredible innovations will occur only when sufficient computing capability is present. For example, no one predicted the first spreadsheet, and until the first PC appeared on the scene, there was no framework in which such an innovation could come about. Our job is to create the microprocessor and PC platform infrastructure with ever-increasing power and capability; innovative ideas for using them will follow.

As mentioned earlier, one area that I believe will require huge numbers of MIPS (not to mention bandwidth) is human interface enrichment: 3D, rich multimedia, sight, sound, and motion. Tomorrow's applications will increasingly incorporate video, sound, animation, color, 3D images, and other visualization techniques to make PCs and applications easier to use.

The system-on-a-chip, microsystems computer industry

Gordon Bell, Microsoft Corp.

The inevitability of complete computer systems on a chip will create a microsystems industry. In addition, forecasters predict 32-Mbyte memory chips by 1999. So by 2002 we would expect a personal computer on a chip with at least 32 Mbytes, video and audio I/O, built-in speech recognition, and industry-standard buses for mass storage, local area network, and communications.

Technology will stimulate a new computer industry for building application-specific computers that require partnerships among system customers, chip fabricators, ECAD suppliers, intellectual property (IP) owners, and systems builders.

The volume of this new microsystems industry will be huge—at least two orders of magnitude more units than the PC industry. For every PC, there will be thousands of other kinds of systems built around a single-chip computer architecture with on-chip interconnection bus. This architecture will be complete with processor, memory hierarchy, I/O (including speech), firmware, and platform software. Powerful processors will enable firmware to replace hardware.

Silicon Graphics (Mips) supplies the key technology for Nintendo and Sony to build games, and WebTV to build an Internet access set-top; Netscape's Navio licenses software to build Internet consumer access devices including phones, games, and television sets that attempt to replace PCs. (Partners included IBM, NEC, Nintendo, Oracle, Sega, and Sony.) Sun's Microelectronics Division is designing and licensing special processors for the Java language and environment. Acorn licenses its ARM processor. Oracle is licensing its network computer to sell server software. Microsoft has various alliances for designing pocket and set-top computers.

The emerging microsystems industry will encompass

- customers building microsystems for embedded applications like automobiles, room and person monitoring, PC radios, PDAs, telephones, set-top boxes, videophones, and smart refrigerators;
- about a dozen foundries that fabricate microsystems—many in Japan and Korea;
- custom companies such as VLSI Technology and LSI Logic that supply "core" IP and take the systems responsibility;
- existing computer system companies like Digital Equipment Corporation, Hewlett-Packard, IBM, Silicon Graphics, and Sun that have large software invest-

ments tied to particular architectures and software;
- fab-less and chipless IP companies that supply designs for royalty;
- ECAD companies that synthesize logic and provide design services (Cadence and Synopsys);
- circuit wizards who design fast or low-power memories (VLSI libraries), analog for audio (which is also a DSP application), radio and TV tuners, cellular radios, GPSs, and micromechanical structures;
- varieties of processors from traditional CISC and RISC to DSP and multimedia;
- computer-related applications that require designers to understand a great deal of software and algorithms (communications protocols and MPEG); and
- proprietary interface companies like Rambus developing proprietary circuits and signaling standards (traditional IP).

Like previous computer generations stemming from Moore's law, a microsystem will most likely have a common architecture. It will consist of an instruction set architecture such as that of the 8088, Mips, or ARM; a physical or bus interconnect that is wholly on the chip and used to interconnect processor memory and a variety of I/O interfaces (disk, Ethernet, audio); and software to support real-time and end-use applications. As in the past, common architectures are essential to support the myriad of new chips economically.

Will this new industry just be an evolution of custom microcontroller and microprocessor suppliers, or a new structure like the one that created the minicomputer, PC, and workstation industries? Will computer companies make the transition to microsystems companies, or will they just be IP players? Who will be the microsystems companies? What's the role for software companies?

Thirty-six ECAD, computer, and semiconductor firms announced an "alliance" for this purpose on September 4, 1996. [See IEEE Micro, Oct. 1996, p. 2.—Ed.]

Gordon Bell is a computer industry consultant-at-large and senior researcher at Microsoft Corp. in Washington, and former head of R&D at Digital. He is a member of various boards, has participated in several start-ups (including the Computer Museum), authored High Tech Ventures, and won various awards including the 1991 National Medal of Technology and the IEEE von Neumann Medal. http://www.research.microsoft.com/research/barc/gbell.

The consumer market segment, rather than the business market segment, is driving PC development in this area. Although the business market struggles with how to interpret and present enormous amounts of information more clearly, home users are leading business people in discovering creative ways to solve problems graphically. There are huge opportunities for enterprising application designers to incorporate 3D visualization in clarifying complex business information. More powerful processors with powerful graphics make it easy to display information visually rather than numerically and therefore easier to interpret the information. PCs with smart user interfaces will enable their users to become

active seekers of information rather than passive absorbers.

Some argue that, in the face of the runaway success of the Internet, less rather than more processing power is needed on the desktop. So-called network computers on the drawing board today allow users to download necessary "applets" and data for temporary use. These devices may find a niche, but the amount of processing power on the desktop (or in the living room) will depend on the kind of Internet experience users wish to have. If they simply want to browse through traditional data types, a less powerful processor may suffice. However, if they want a rich multimedia experience, viewing information with 3D images and sound will require considerable MIPS.

Another area that urgently needs attention is the historic lag between hardware and software development. Software has always lagged behind available hardware; just as an application takes advantage of new hardware capabilities, vendors release the next generation of hardware. Widespread object-oriented design may help close this gap, but we need breakthroughs in software development to help software keep pace with hardware developments. I believe this is an area of enormous opportunity. Whoever is first to fully take advantage of the coming microprocessor power to offer innovative applications will be the unquestioned leader.

THE MICROPROCESSOR DEVELOPMENT path we've been on for the past 25 years can easily continue into the next 10. Performance can continue to advance until we reach close to a stunning 400 million transistors on a 1.7-inch chip in the year 2006. However, manufacturing capital costs will be in the multibillion-dollar range, necessitating huge volumes to drive down unit price. Besides the huge cost of manufacturing, we have big technological hurdles to overcome before we realize such a chip. We need to know how to test and validate 400 million transistors, how to connect them, power them, and cool them.

Once in hand, however, computing power of such magnitude will set the stage for huge innovations and market segment opportunities in everything from business computing to "edu-tainment" products for kids. One thing I can predict with certainty: Micro 2006 will surprise us all with applications and devices that will dramatically change our world. 🔲

Acknowledgments

I thank fellow Intel employees Richard Wirt and Wen-Hann Wang for assistance in gathering and formulating prediction data.

References

1. P.P. Gelsinger et al., "Microprocessors Circa 2000," *IEEE Spectrum*, Oct. 1989, pp. 43-47.
2. P.P. Gelsinger et al., "2001: A Microprocessor Odyssey," *Technology 2001, The Future of Computing and Communications*, D. Leebaert, ed., The MIT Press, Cambridge, Mass., 1991, pp. 95-113.
3. *The National Technology Roadmap for Semiconductors*, Semiconductor Industry Assoc., San Jose, Calif., 1995.
4. R.P. Colwell and R.L. Steck, "A 0.6 μm BiCMOS Processor With Dynamic Execution," *Proc. Int'l Solid-State Circuits Conf.*, IEEE, Piscataway, N.J., 1995, p. 136.
5. U. Weiser, "Intel MMX Technology—An Overview," *Proc. Hot Chips Symp.*, Aug. 1996, p. 142.
6. "Special Issue: Celebrating the 25th Anniversary of the Microprocessor," *Microprocessor Report,* Aug. 5, 1996.
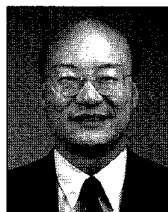
Albert Y.C. Yu is senior vice president and general manager of the Microprocessor Products Group at Intel Corporation. He has responsibility over Intel Architecture Processor products such as the Pentium, Pentium Pro, and future microprocessors. He also oversees platform architecture, design technology, microprocessor software products, and microcomputer research labs.

Yu received his PhD and MS from Stanford University and his BS from the California Institute of Technology, all in electrical engineering. He is a senior member of the IEEE and the Computer Society.

Address questions or comments about this article to Albert Y.C. Yu, Intel Corporation, M/S RN 3-31, 2200 Mission College Blvd., San Jose, CA 95052-8119; Albert_Yu@ ccm.sc.intel.com.

Reader Interest Survey

Indicate your interest in this article by circling the appropriate number on the Reader Service Card.

Low 162   Medium 163   High 164