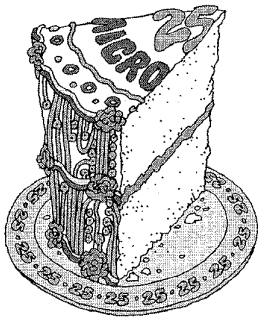
THE MICROPROCESSOR TODAY

Michael Slater

MicroDesign Resources



Microprocessor Report*'s publisher outlines technology and business issues in today's microprocessor industry.*

From their humble beginnings 25 years ago, microprocessors have proliferated into an astounding range of chips, powering devices ranging from telephones to supercomputers. Today, microprocessors for personal computers get widespread attention—and have enabled Intel to become the world's largest semiconductor maker. In addition, embedded microprocessors are at the heart of a diverse range of devices that have become staples of consumers worldwide.

Microprocessors have become specialized in many ways. Those for desktop computers fall into classes based on their instruction set architectures: either x86, the primary surviving complex instruction set computing (CISC) architecture, or one of the five major reduced instruction set computing (RISC) architectures—PA-RISC, Mips, Sparc, Alpha, and PowerPC. Such chips typically integrate few functions other than cache memory and bus interfaces with the processor but usually include a floating-point unit and memory management unit.

Embedded microprocessors, on the other hand, typically do not have floating-point or memory management units but often integrate various peripheral functions with the processor to reduce system cost. This makes them more application specific, leading to a massive proliferation of devices characterized not only by their processor's instruction set and core CPU performance but also by their on-chip peripherals.

Digital signal processors (DSPs) are the most specialized embedded microprocessors. Designed for real-time processing of digitized analog signals, these processors have unique instruction sets and other architectural features that give them high performance for a relatively narrow range of tasks. Recently, a new class of DSPs, called media processors, has emerged to handle audio, video, graphics, and communication tasks in multimedia PCs.

Although the desktop computer market tends to discard old processors in just a few years, many processors survive for an amazingly long time in the embedded market. Personal computers have moved from 8- to 16-bit and now to 32-bit processors, and many workstations and servers are already using 64-bit microprocessors. In the embedded market, however, even 4-bit microprocessors continue to sell well, and 8-bit devices lead in volume.

Figure 1 shows the changes in market share for 32- and 64-bit microprocessors over the past five years. Driven by the success of the PC, the x86 architecture has dominated. Motorola's 68000 has held a strong second-place position in units (though not in dollars), with the embedded market as its stronghold. Hitachi's SuperH has come from nowhere to take third place, while other architectures have much more modest positions. As the dramatic changes during this period show, market share can be quite turbulent among these second-tier embedded processors.

There are far more microprocessors available for sale today than it would be possible to describe, even briefly, in an article such as this. The current market includes more than 50 surviving instruction set architectures, hundreds of different implementations, and thousands of minor variations. Rather than attempting to be comprehensive, I will focus on a selection of the leading-edge microprocessors and issues in each market segment.

Dividing up the market

Microprocessors for personal computers get the most public attention because the performance and compatibility of PCs depend on the microprocessors at their cores. In recent years, PC microprocessors have become so high profile that a bug of minor significance in Intel's Pentium proces-

0272-1732/96/$5.00 © 1996 IEEE

sor captured headlines in the mainstream press for months. Microprocessor commercials even appear on prime-time television. Embedded applications, on the other hand, have a fixed complement of software, so the microprocessor inside is of relatively little interest to the consumer.

PC microprocessors are also a major profit area in the microprocessor business. PCs are the only application so far that uses expensive microprocessors—typically costing $75 to $500—in volumes of tens of millions per year. Embedded applications use far more processors—literally billions per year—but most of them are very low cost devices, selling for under $5, with thin profit margins. Some embedded applications use processors in the same price range as PCs, but they are very low volume—often only thousands of units per year. Today, even among 32-bit embedded processors, more than 60 chips sell for less than $40 and some for less than $10.

Microprocessors for PCs generate far more profit than embedded processors for two reasons. First, because the PC is the most expensive device to use a microprocessor in high volume, PC makers can afford more expensive processors. When coupled with large amounts of DRAM, disk drives, and CRT displays that cost hundreds of dollars each, even a relatively expensive microprocessor does not dominate system cost. Second, the overwhelming importance of software compatibility in the PC market has enabled Intel to achieve tremendous control over the PC microprocessor market. With no close competitors, Intel can enforce much higher profit margins than makers of embedded microprocessors, which must compete in a field where instruction set compatibility commands less value.

Unix workstations constitute the high-performance segment of the desktop computer market. (I use the term desktop computer as a shorthand for a single-user, general-purpose computer; in this context, it includes deskside and portable systems as well as true desktops, workstations as well as PCs, and even many servers.) The workstation market has some of the same attributes as the PC market, in that application software compatibility with previous-generation systems is of great importance. There are two big differences, however. First, the total Unix workstation market, in units, is less than 1% of the PC market, and second, performance is more important than price. As a result, microprocessors for workstations are typically designed with performance as a higher priority than price.

This digression into business issues is necessary for any comprehensive discussion of the microprocessor industry, because these issues have a pervasive effect on microprocessor design and manufacturing. Companies such as Intel and Motorola fund leading-edge manufacturing plants for logic devices with the high profits they make (or hope to make) supplying microprocessors for desktop computers. In contrast, makers of embedded processors generally get by with lower cost manufacturing processes and depreciated fabrication plants. The divergent characters of the embedded and PC microprocessor markets also drives the evolution of instruction set architectures: New architectures find relatively easy entry into the embedded market, while desktop processor makers go to great lengths to stick with dominant architectures.

Economics have had a huge effect on microprocessors designed for workstations. Because they are built in much lower volumes than PC microprocessors, their amortized design cost per unit is much greater. Such processors have survived primarily because they enable system businesses that produce far more revenue (and profit) than would the sale of a comparable number of microprocessors. This economic reality has led to large workstation makers designing nearly all of today's workstation-specific microprocessors, with chip companies playing little more than a foundry role.

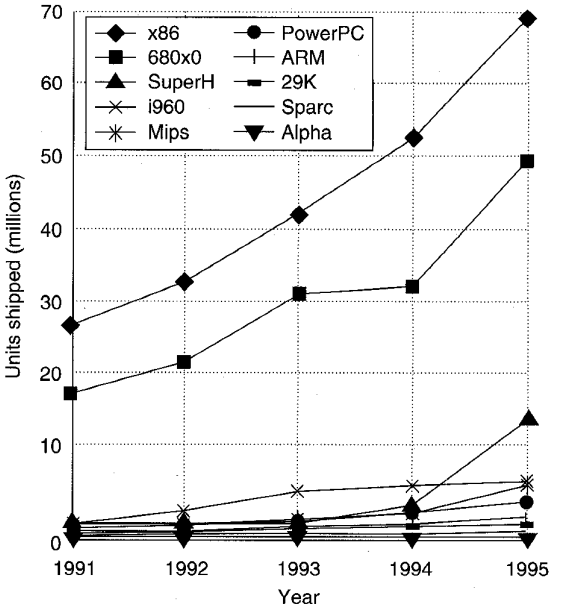


Figure 1. Unit shipments of leading 32- and 64-bit architectures.

As microprocessors for PCs have increased in speed, they have reached a performance level that makes them suitable for much of the workstation market. This factor, combined with the emergence of Microsoft's Windows NT as a viable alternative to Unix as a workstation operating system, is leading to a convergence of the technology for PCs with that for workstations. Ultimately, this convergence will threaten the survival of some of the workstation processor architectures.

The battle for the desktop

Two operating systems account for the vast majority of desktop computer use today: Microsoft Windows (in its various versions) and Macintosh. IBM's OS/2 is in third place, and Unix is a distant fourth. Windows' popularity has given Intel's x86 architecture a preeminent position, while the Macintosh has established PowerPC as the only RISC architecture with a significant share of the desktop market. Even so, the Mac's share is modest—around 8%—and shows no signs of increasing.

Nearly every maker of RISC microprocessors has dreamed of capturing part of the x86 architecture's high-volume market. The first to try was Sun, which hoped to repeat the PC industry phenomenon by making its Sparc-based workstations, running Sun's version of Unix, an openly licensed stan-

Table 1. Originators and licensees for leading desktop architectures.

Architecture	Originator	Licensees
Alpha	Digital Semiconductor	Mitsubishi, Samsung
Mips	Mips Technologies	IDT, NEC, Toshiba
PA-RISC	Hewlett-Packard	Hitachi, Samsung
PowerPC	Apple, IBM, Motorola	Groupe Bull, Exponential
Sparc	Sun Microelectronics	Fujitsu (includes HaL, ICL, and Ross)

dard. This effort failed miserably for at least two reasons. First, Unix was, and is, unsuitable for a mass-market operating system because of its complexity, its resource requirements, and its lack of personal productivity applications. Second, Sun was never willing to let Sun-compatible system makers operate unrestrained, for fear of the effect that might have on Sun's own hardware business.

Next up was Mips, then an independent company (Mips Computer Systems). Microsoft chose Mips as the first RISC architecture that the emerging high-end version of Windows, Windows NT, would support. Mips engineered the ACE initiative, and at one point had both Digital Equipment Corporation and Compaq planning to build Mips-based systems to run Windows NT. But the timing was bad; Compaq fell on hard times as lower cost PC clones eroded its business, and Digital decided to create its own architecture and abandon Mips.

Mips made its own contribution to the failure of ACE by trying to collect large license fees for reference system designs. Silicon Graphics (SGI) soon thereafter swallowed Mips Computer Systems, which became SGI's Mips Technologies subsidiary. SGI has shown no interest in either Windows NT or the high-volume desktop market, and Microsoft is dropping Mips support in Windows NT.

Then came PowerPC, backed by IBM, Motorola, and Apple. Apple has successfully converted the Macintosh line to PowerPC, giving PowerPC the biggest desktop market of any RISC processor. Because Apple and the Macintosh platform itself are struggling to maintain their modest position, however, this doesn't represent much of a growth opportunity. Efforts to go beyond Macintosh on the mainstream desktop have largely failed: IBM's OS/2 for PowerPC was stillborn, Taligent folded its tent, and PowerPC's position in the Windows NT market is weak.

The most recent architecture to aim at the desktop market is Alpha, Digital's home-brewed replacement for the Mips architecture. Digital wholeheartedly embraced Windows NT and has the benefit of owning its own systems business, including a PC business. But Windows NT is only now approaching the maturity that will enable it to become a mainstream operating system; and DRAM prices are only now becoming low enough to render Windows NT's additional memory requirement insignificant.

Until now, Digital has used Alpha's outstanding performance to sell very fast systems at premium prices—a nice niche business, but hardly a factor in the PC market. Next year, Digital plans to begin moving down into high-end PCs, setting the stage for an eventual attack on the mainstream PC market. It is a long shot for Alpha to capture a significant mainstream role, but at least it can't be counted out yet.

It is in this light that PowerPC's position in the Windows NT market appears so weak. PowerPC processors are not nearly as fast as the Alpha chips and don't offer a significant performance advantage over Intel processors, leaving them between a rock and a hard place. Customers looking for safety and compatibility choose Intel; those seeking maximum performance on a small set of applications are drawn to Alpha. This leaves few for whom PowerPC would be a compelling choice.

The staying power of the PowerPC backers is the architecture's key strength. If a future generation of chips is much stronger than today's, the architecture could end up head to head with Alpha in an attempt to capture the number two position in the Windows NT market.

This leaves Hewlett-Packard's PA-RISC as the only RISC architecture whose owner never attempted to use it in an attack on Intel's market share. This may have been an excellent decision, considering the fate of companies that have tried. HP is now engaged with Intel in a joint development project that will lead to a new architecture around 1998. The architecture, called IA-64 and to be first implemented in a chip code-named Merced, will provide backward compatibility with both x86 and PA-RISC programs. Having built a large computer business around its architecture, HP has found no compelling reason to spend billions of dollars on fabrication facilities and chip designs to provide processors for these systems. Thus, it has joined future paths with Intel.

Table 1 shows the companies backing each of the architectures for general-purpose computers. (Not shown are licensed implementations; for example, Cyrix has licensed its x86 processor designs to IBM Microelectronics and SGS-Thomson.) Although there has been a mad rush to sign up licensees, it has turned out to be relatively insignificant; the owners and primary backers of each architecture determine its fate.

Pentium dominates computing today

Intel's Pentium processor series dominates today's desktop computer market. Depending on clock speed, this chip spans a price range (in quantities of 1,000) from about $75 to just over $500, putting it at the appropriate price points for most PCs. Although early Pentium processors provided little advantage over 486 chips, Intel's aggressive promotion of Pentium and rapid increase in the chip's clock speed enabled it to sweep the desktop market by the end of 1995 and the notebook market in the first half of 1996.

Following a familiar pattern in the microprocessor industry—but at an accelerated pace—Intel has twice moved Pentium to a new process technology. The initial chips, code-named P5, were built in 0.8-micron BiCMOS and ran at 60 and 66 MHz. These chips were power hungry, and Intel phased them out before Pentium began its move into the mainstream PC market. The next version, the P54C, shrank the design to 0.6-micron BiCMOS and enabled clock rates of 75 to 120 MHz. This version also cut the supply voltage to 3.3 V and added dynamic power management circuitry.

This feature shuts down portions of the chip not in use on a cycle-by-cycle basis, slashing typical power consumption. Then Intel shrank the design once again to 0.35-micron BiCMOS, enabling clock speeds up to 166 MHz. A minor revision of this design pushes the clock speed to 200 MHz—more than three times that of the original Pentium.

To keep system design relatively easy, however, Intel has held the system bus speed at 60 or 66 MHz. Because of this, there is a huge gap between increasing core CPU speeds and the bandwidth of the external bus, which provides access to the level-two cache as well as to main memory. This reduces the benefit of faster core speeds; the 200-MHz Pentium has a typical performance gain of less than 10% over the 166-MHz chip. Power consumption has also crept up to uncomfortable levels as the clock speed has increased, keeping the 166-MHz and faster chips out of portable systems.

Intel will mitigate these problems early next year with a new version of Pentium, code-named the P55C and implemented in 0.28-micron CMOS. By doubling the size of the on-chip cache, Intel estimates that the miss rate will decrease 20 to 40% on typical Windows applications, mitigating the performance loss from the relatively slow external bus. The P55C will also include pipeline enhancements to boost its per-clock performance, as well as the MMX instruction set extensions for multimedia (described later).

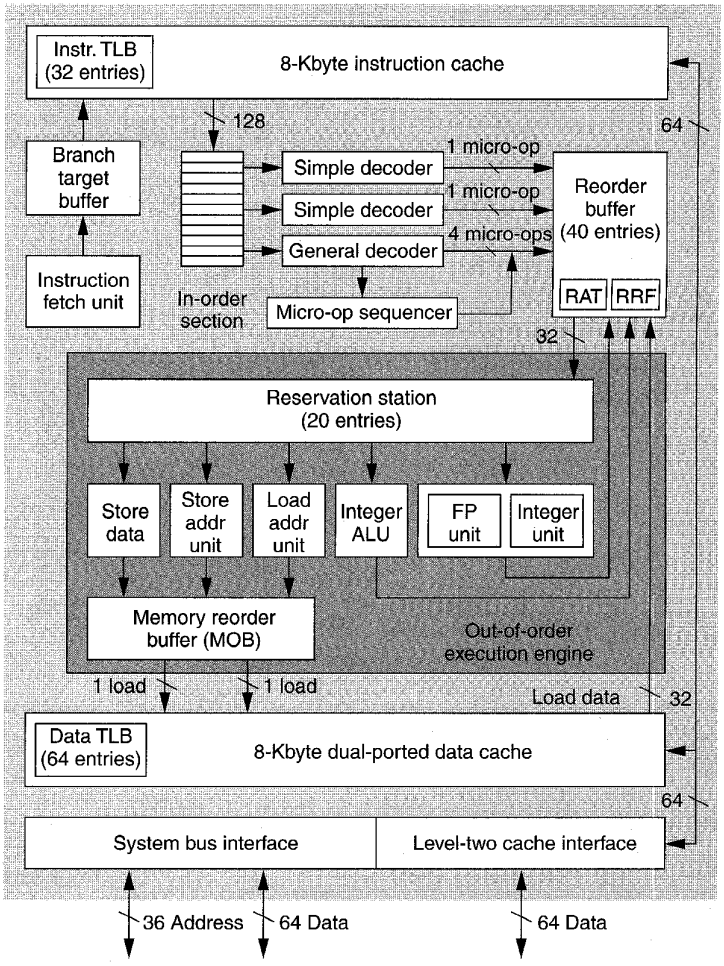


Figure 2. Pentium Pro microprocessor block diagram. TLB: translation look-aside buffer.

The P55C will mark Intel's shift away from the BiCMOS process technology of earlier Pentiums. The 0.28-micron (drawn gate size) process enables Intel to reduce the supply voltage from 3.3 to 2.8 V, which significantly reduces power consumption. At this low voltage, however, bipolar transistors offer little benefit, making the extra process steps of BiCMOS unjustified. The supply voltage reduction will make higher clock rates practical for portable systems and will simplify cooling in desktop systems.

Intel's new frontier: Pentium Pro

The Pentium design uses a simple, restrictive approach to superscalar operation. Its two pipelines do not operate entirely independently; when one stalls, the other must stop as well, so no out-of-order execution is allowed. Furthermore, the floating-point unit is not autonomous but relies on the integer pipelines, so integer and floating-point instructions cannot execute in parallel.

Intel's most recent microprocessor design, Pentium Pro (P6), takes a far more aggressive approach to deliver more performance per clock cycle while also enabling higher clock speeds. Figure 2 shows the processor's block diagram.

The Pentium Pro design completely decouples instruction dispatch and execution, translating x86 instructions into internal micro-operations, not unlike traditional microcode instructions. These micro-ops then pass to a 40-entry reorder buffer, where they are stored until any required operands are available. From there, they are issued to a 20-entry reservation station, which queues them until the needed execution unit is free. This design allows micro-ops to execute out of order, making it easier to keep parallel execution resources busy. At the same time, the fixed-length micro-ops are easier to handle in the speculative, out-of-order core than complex, variable-length x86 instructions.

To enable high clock speeds, Pentium Pro is very deeply pipelined (also called superpipelined). Because the reservation station represents an elastic element, the pipeline does

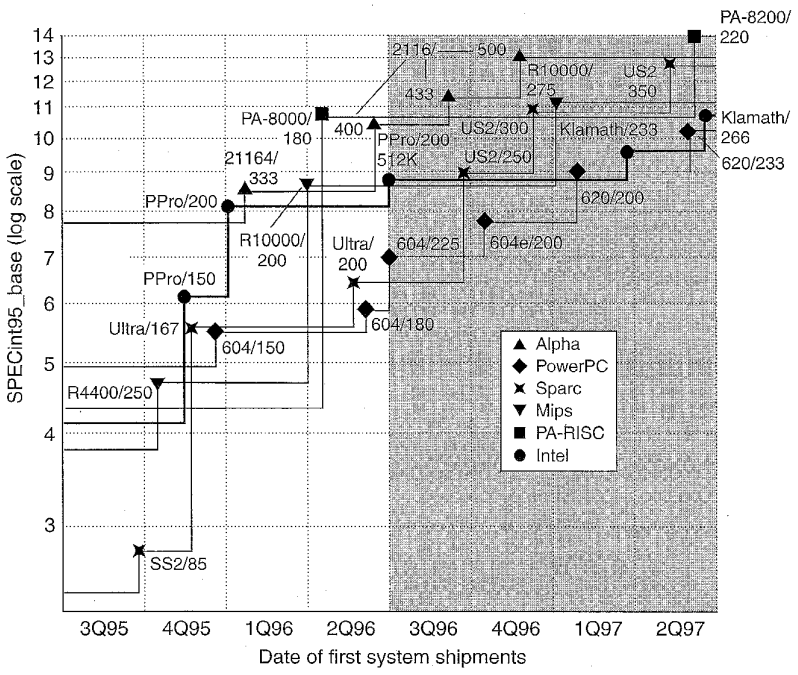


Figure 3. SPECint95 (base) performance versus time for x86 and RISC architectures. Numbers following processor names are clock speeds in MHz.

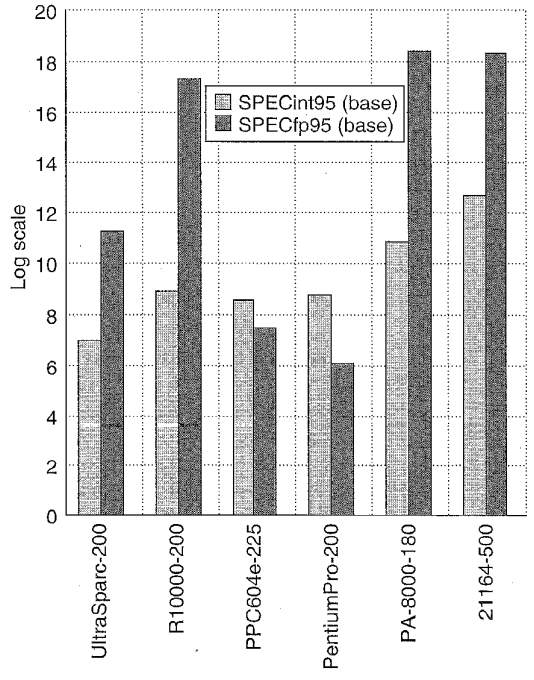


Figure 4. SPEC (base) microprocessor performance for high-end microprocessors shipping today.

not have a fixed number of stages, but the minimum number of clock cycles for an instruction to complete is 12. Cache access and instruction decoding are each split across two and one-half clock cycles.

To push Pentium Pro performance as high as possible, Intel designed a special level-two cache chip that is mounted in the same package with the CPU chip. The connections between the CPU and the cache chip are point to point and don't leave the package, which enables Intel to use nonstandard voltage levels and achieve high data rates. The level-two cache chip, which Intel makes in both 256- and 512-Kbyte versions, delivers 64 bits per clock cycle, even with CPU clock speeds up to 200 MHz.

This cache strategy was effective in bringing Pentium Pro to market with performance numbers that sent shock waves through the planning departments of most RISC microprocessors makers. As Figure 3 shows, at its introduction Pentium Pro exceeded the SPECint95 performance of all shipping RISC microprocessors. This position didn't last long, however, as Intel has gone more than a year without either increasing clock speed or introducing a new microarchitecture. Each of Intel's RISC competitors has done one or both. As Figure 4 shows, Pentium Pro is even further behind the RISCs when it comes to floating-point performance.

In the long run, however, Intel doesn't want to devote half its fab capacity to relatively low-margin SRAMs; it has been working with SRAM makers to provide industry-standard memory chips for future versions of the P6. In particular, Intel has disclosed plans for the P6 series' second member, code-named Klamath, which will use external SRAMs running at half the processor's clock speed for the level-two cache and will implement the MMX instruction set extensions. Intel intends Klamath, due in the first half of 1997, to be the P6 chip that drives that architecture into mainstream systems.

Going after Intel directly

RISC microprocessor makers have tried to take some of Intel's market share by leveraging their superior instruction sets to produce faster and less expensive processors. This has proven a very difficult game, however, due primarily to the enormous software barriers that new architectures in the desktop market face.

Other companies have challenged Intel on its own turf, building microprocessors that run the same software as Intel's chips. Today, the primary players are Advanced Micro Devices (AMD) and Cyrix Corporation, along with Cyrix licensees IBM Microelectronics and SGS-Thomson. Texas Instruments also serves the low end of the market with 486

Table 2. Key features of selected x86 microprocessors.

	Intel						AMD		Cyrix	
Feature	Pentium P54C	Pentium P54CS	Pentium P55C	Pentium Pro P6	Pentium Pro P6S	Klamath	K5	K6	6x86 M1	M2
Maximum clock rate (MHz)	120	200	200	150	200	233-266†	100	>180	150	225
Pinout	P54C	P54C	P54C	PPro	PPro	Klamath	P54C	P55C	P54C	P55C
Cache (data/ instr., Kbytes)	8/8	8/8	16/16	8/8	8/8	N.A.	8/16	32/32	16/16	64 (unified)
MMX	No	No	Yes	No	No	Yes	No	Yes	No	Yes
Decode rate (instr./clock cycle)	2	2	2	3	3	3	1-4	2	2	2
Issue rate per clock cycle	2 instr.*	2 instr.	2 instr.	5 micro-ops	5 micro-ops	5 micro-ops	4 micro-ops	4 micro-ops	2 instr.	2 instr.
Out-of-order execution	No	No	No	Yes	Yes	Yes	Yes	Yes	Limited	Limited
Die size (mm^2)	148	90	140	308	196	N.A.**	181	~180	167	<200
Transistors (millions)	3.3	3.3	4.5	5.5	5.5	N.A.	4.0	8.8	3.3	6.0
Process (µm/layers)	0.5/4 BiCMOS	0.35/4 BiCMOS	0.28/4 CMOS	0.5/4 BiCMOS	0.35/4 BiCMOS	0.28/4 CMOS	0.35/3 CMOS	0.35/5 CMOS	0.44/5 CMOS	0.35/5 CMOS
Mfg. cost†	$50	$40	$60	$180††	$145††	N.A.	$70	$85	$70	$95
Production	Now	Now	1Q97	Now	Now	1H97	Now	1H97	Now	1H97
List price***	$106-134	$204-509	N.A.	$534	$428-1,035	N.A.	$60-134	N.A.	$98-299	N.A

* Indicates x86 instructions
** Not available
*** As of 4Q96, quantities of 1,000
† MicroDesign Resources estimates
†† Includes 256-Kbyte level-two cache

microprocessors. Table 2 summarizes the key features of today's most important x86 microprocessors.

AMD has a long history as an alternative supplier of x86 microprocessors. The company was a licensed alternate source of Intel's 8086 and 286 microprocessors, but the technology exchange agreement between the two companies broke down into a bitter and drawn-out arbitration. As a result, Intel never transferred its 386 or later technology to AMD. Instead, AMD entered the 386 and 486 markets by reverse-engineering Intel's chips. This involved extracting the circuit designs, making minor modifications (such as for static rather than dynamic operation), and producing new physical layouts tuned for AMD's process technology.

This path proved successful in that it enabled AMD to continue supplying microprocessors to the PC industry. However, it offered AMD little opportunity for differentiation and no chance of catching up with Intel's performance level. AMD couldn't even begin its reverse-engineering and reimplementation process until Intel shipped a product.

AMD therefore decided to create an entirely independent design, taking from the Intel chips only the instruction set (for software compatibility) and the bus interface and pinout (for system interface compatibility). After several delays, the K5 reached the market, but without delivering the anticipated performance level. The chip was supposed to deliver performance 30% higher than an Intel Pentium processor at the same clock rate. Instead, it barely matched Intel's per-clock performance on Windows application benchmarks, despite a much more complex design and a 30% greater transistor count.

As of October 1996, AMD had been unable to make the chip run faster than 100 MHz, while Intel was shipping Pentiums at up to 200 MHz. This failing relegated AMD to the low end of the PC microprocessor business, leaving little profit for a chip as large as the K5 (see Table 2). At the same time, the 486 market had largely dried up, and what remained was priced in the $20-30 range, leaving AMD no significant older products to fall back on.

AMD recently released an improved version of the K5 design that eliminates bottlenecks and reaches the originally targeted performance levels. At 100 MHz, it delivers performance equivalent to a 133-MHz Pentium, moving AMD into the midrange Pentium market.

AMD's big opportunity, however, depends on the K6—a design that started life as the NexGen 686, which AMD bought NexGen to obtain. Like the Pentium Pro and K5, the K6 uses a decoupled decode/execute design in which x86 instructions are first decoded into internal, RISC-like operations. AMD also is adding the MMX instruction set extensions. As the K5 design has shown, though, the devil is in the details: A design's effectiveness depends on a multitude of subtle design issues, any one of which can become a performance-limiting bottleneck. On paper, the K6 looks good, but until AMD ships its first K6 samples, due by the end of 1996, how well it performs will remain an unknown.

Unlike AMD, Cyrix designed its own x86 cores from the start. The company started with a low-end 486-class core,

Table 3. Key features of selected high-performance microprocessors. ***(Source: Vendors except where noted)***

Feature	Digital 21164	PowerPC 620	PowerPC 604e	Sun UltraSparc	Micro Sparc-2	HP PA-8000	HP PA-7300LC	Mips R10000	Mips R5000	Pentium Pro
Clock rate (MHz)	500	200	225	250	110	180	160*	200	180	200
Cache size (Kbytes)	8/8/96	32/32	32/32	16/16	16/8	None	64/64	32/32	32/32	8/8
Issue rate (instr./cycle)	4	4	4	4	2	4	2	4	2	3
Pipeline stages	7	5	6	6/9	5	7-9	5	5-7	5	12-14
Out-of-order execution	6 loads	16 instr.	16 instr.	None	None	56 instr.	None	32 instr.	None	40 ROPs
Rename registers	None 8 FP	8 int/ 8 FP	12 int/	None	None	56 total	None	32int/ 32FP	None	40 total
Memory bandwidth (Mbytes/s)	~400	1,200	~180	1,300	~100	768	213	539	~160	528
Package, pins	CPGA-499	CBGA-625	CBGA-255	PBGA-521	CPGA-321	LGA-1,085	CPGA-464	CPGA-527	SBGA-272	MCM-387
Process (µm/layers)	0.35/4	0.35/4	0.35/4	0.29/5	0.4/3	0.5/4	0.5/4	0.35/4	0.35/3	0.35/4
Die size (mm^2)	209	240*	148	149	233	345	259	298	84	196
Transistors (millions)	9.3	6.9	5.1	3.8	2.3	3.9	9.2	5.9	3.6	5.5
Estimated mfg. cost*	$150	$210	$60	$90	$80	$290	$95	$160	$25	$175**
Maximum power (W)	25	30	20*	30	9	>40	15	30	10	35**
SPEC95 baseline performance (integer/FP)	12.6/18.3	9.0/9.0*	8.5/7.0	8.5/15	1.4/1.9	10.8/18.3	5.5/7.3	8.9/17.2	4.0/3.7	8.7/6.0
Availability	Now	1H97	Now	Now	Now	Now	Now	Now	Now	Now
List price (1,000)	N.A.	N.A.	$594	$1,995	$379	N.A.	N.A.	$3,000	$365	$1,035

* MicroDesign Resources estimate ** Includes 512-Kbyte level-two cache

which it leveraged into a range of products from the 386SX-pin-compatible 486SLC to a 486DX2. Cyrix abandoned these products at the end of 1995, however, as it began the switch to its Pentium-class core, code-named the M1 and officially called the 6x86. This chip delivers impressive performance per clock cycle: At 133 MHz, for example, it outperforms a 166-MHz Pentium on common Windows application benchmarks. Rather than using the complex decoupled decode/execute approach of Pentium Pro and the K5, the 6x86 extends Pentium's relatively straightforward dual-pipeline approach with additional features that enable both pipelines to run concurrently more often.

If Cyrix had access to Intel's leading-edge process technology, its chips might match Intel's Pentium clock rates. But as things stand, Cyrix uses 0.44-micron CMOS technology to compete against Intel's 0.35-micron chips. That Cyrix can beat Intel's Pentium performance even with this handicap is a testament to the efficiency of its design.

Like AMD, Cyrix will move to a next-generation design in early 1997 that will be key to its future success. Code-named the M2, this chip is based on the 6x86 core but adds a much larger 64-Kbyte cache and other performance enhancements, as well as the MMX instruction set extensions.

In 1997, makers of leading-edge PCs will be able to use Intel's P55C or P6-series chips, AMD's K6, or Cyrix's M2. Intel is all but guaranteed the lion's share of the market, but AMD and Cyrix have the opportunity to gain a minority share big enough to be quite significant for them—if they execute well.

By the end of 1997, however, there may be other competitors to contend with. Texas Instruments has a long-pending effort to develop its own x86 CPU core; at least four start-ups in the United States are working on x86 microprocessors; and semiconductor makers in Korea and Japan are probably considering similar efforts as well.

The pursuit of speed

In the never-ending pursuit of maximum performance, microprocessor makers have followed a variety of strategies. In each case, designers must make countless judgment calls—generally backed by simulations—on myriad design options, hoping to make the best use of transistor budgets. Table 3 summarizes the key characteristics of today's highest performance microprocessors.

Perhaps the most fundamental trade-off is between doing lots of work in each clock cycle—which tends to generate complex designs with limited clock rates—or streamlining the design as much as possible in pursuit of maximum clock speed. Sun's SuperSparc is a notable example of a chip that pushed complexity too far, giving up too much in clock rate to justify the per-clock efficiency. Sun remedied this in its

next design, UltraSparc.

Digital has been the most successful proponent of the maximum clock speed approach. The company plans to ship 500-MHz processors this year, while most other vendors' chips will be at 200 to 250 MHz. Digital's Alpha 21164 does deliver the industry's best performance, but not by as big a margin as the high clock speed would indicate. As part of the speed/complexity trade-off, it has among the industry's worst performance per clock cycle. Figure 5 shows a block diagram of the 21164.

Cache strategy is another area where many approaches are possible. Here again, Digital stands out from the pack, with the only microprocessor with a two-level cache on chip. Separate 8-Kbyte first-level instruction and data caches enable single-cycle access even at high clock rates. A slower 96-Kbyte second-level cache provides faster access than could an external cache. Typical 21164 system designs have an external level-three cache.

Intel's Pentium Pro has the smallest on-chip caches of any high-performance processor: a mere 8 Kbytes each for instructions and data. This is because the custom-designed level-two cache chip (described earlier) is mounted in the same package as the processor and can deliver near on-chip speeds. Most other high-performance processors have on-chip level-one caches of either 16 or 32 Kbytes each for instructions and data. HP's 7300LC has the largest caches, at 64 Kbytes each.

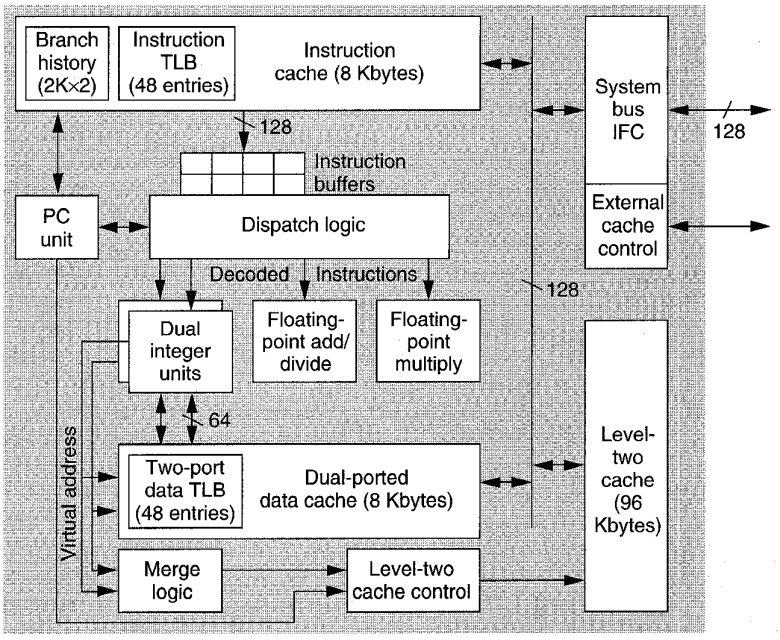


Figure 5. Block diagram of Digital's Alpha 21164 microprocessor.

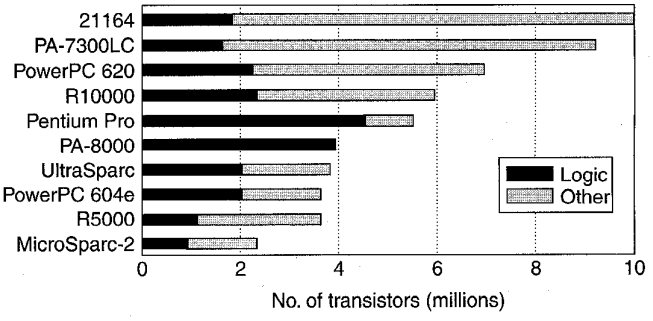


Figure 6. Allocation of transistors in high-end microprocessors.

All of today's high-performance microprocessors are superscalar; most issue four instructions per clock cycle. One exception is Intel's Pentium Pro. Its x86 instruction set encodes more functions into each instruction, so it reaches comparable performance levels (on integer code) decoding only three instructions per clock cycle. Pentium Pro must deal with the additional complexity of the x86's variable-length instructions, which make parallel decoding considerably more challenging. It also has an additional block of logic to convert complex x86 instructions into multiple internal instructions. RISC architectures make no such distinction between external and internal instructions.

Most high-performance microprocessors support some degree of out-of-order execution to keep the entire machine from stalling when one instruction stalls. Digital's 21164 designers, in pursuit of high clock speed, provided minimal out-of-order support, allowing reordering of load operations only. Sun likewise avoided the complexity of out-of-order operation. Others allow from 16 to 56 instructions to execute out of order.

Other microarchitectural features for which designers have chosen different strategies include the size of translation look-aside buffers, the complexity of the branch prediction algorithm, and the size of the branch history table. As for external interfaces, 64- or 128-bit-wide data buses are universal, and the fastest devices tend to provide a dedicated bus connected to an external level-two cache.

Figure 6 shows how today's leading high-performance microprocessors allocate transistors. Digital's 21164, with nearly 10 million transistors, is the biggest chip overall, but Pentium Pro, with its more complex instruction set, has more

than twice as many logic transistors. The figure shows that high-end processors today typically have CPU cores with 2 to 4 million transistors devoted to logic. The number of transistors devoted to memory ranges from less than 1 million to more than 6 million.

Extending instruction sets for multimedia

Although the gulf in instruction set design style between the x86 and RISC camps remains, they do agree on one point: Modest extensions to the instruction set can significantly improve multimedia performance. A small increase in die area delivers a significant boost in performance for functions such as MPEG encoding and decoding, audio synthesis, image processing, and modems.

At the heart of most vendors' multimedia extensions are single-instruction, multiple-data (SIMD) operations. By taking a 64-bit ALU and allowing the carry chain to be broken at various points, essentially the same amount of logic can perform two 32-bit operations, four 16-bit operations, or eight 8-bit operations, all in parallel. One complication is that multiple carry bits are not available. Fortunately, however, most signal-processing operations benefit from saturation arithmetic. Instead of rolling over and setting the carry bit, saturation arithmetic sets the result at the minimum or maximum value. Most multimedia extensions add saturation arithmetic as an option. Other common additions are instructions for multiply-add and data element packing and unpacking.

HP was the first to add such extensions to its RISC architecture, but HP's instructions are quite simple. Sun offers the most comprehensive set of extensions in its VIS (Visual Instruction Set), implemented in UltraSparc. Sun's extensions include some relatively complex instructions, such as pixel distance, in addition to the simpler SIMD operations.

The most widely discussed, though not yet shipped, set of extensions is Intel's MMX, which will appear next year in the P55C and Klamath processors. Both AMD and Cyrix will offer MMX-compatible extensions next year as well. Intel estimates that the performance of MMX-enhanced code will be from 1.4 times better for MPEG video decoding to more than 4 times better for still-image processing (such as Adobe Photoshop filtering). Of course, most programs won't benefit at all, and compilers don't use MMX—programmers must handcraft the code to realize the benefits.

The Mips and Alpha camps recently announced their own multimedia extensions, leaving PowerPC as the only popular architecture not to follow suit. This is ironic, since PowerPC's primary user—Apple—focuses on multimedia, and one of the PowerPC's predecessors—Motorola's ill-fated 88110—had a set of graphics instruction set extensions.

Media processors enter the fray

General-purpose microprocessors can improve their handling of multimedia data types through instruction set extensions, but there are compelling reasons to use a separate processor for these tasks. DSP-like architectures provide multiple operand data paths, very-long-instruction-word-like arrangements, and other special features that make them fast but often hard to program. With these characteristics, a given silicon area can deliver much greater performance on signal-processing applications than could an equal area in an extended general-purpose architecture.

DSP chips are not new; indeed, they are at the heart of most modems, cellular phones, disk drives, and countless other devices. They have had little success in PCs, however, because they aren't well optimized for the PC environment. However, several companies are now making media processors carefully designed for PCs. These chips typically have PCI bus interfaces, integrated codecs or codec interfaces, and graphics engines that provide compatibility with legacy PC display controller standards (such as VGA). Most importantly, makers of these PC media processors also provide driver software that enables applications to communicate with the chips via Microsoft's DirectX application programming interfaces (APIs). Thus, programmers need not customize application programs for each hardware design.

Today, a start-up company called Chromatic Research (Sunnyvale, Calif.) is the closest to shipping such a media processor. Like many pioneering microprocessor companies of recent times, Chromatic Research is fabless. LG Semicon and Toshiba manufacture and sell the chips, while Chromatic sells the software that makes them work. Chromatic's Mpact media processor can perform not only 2D and 3D graphics rendering but also MPEG-1 and MPEG-2 decompression, MPEG-1 compression, teleconferencing, 33-Kbps fax/modem, and audio synthesis. Philips has its own media processor, TriMedia; Samsung, Mitsubishi, IBM, and others have media processors in the works.

Whether these media processors have a long-term role in PCs remains a subject of controversy. From Intel's perspective, there is room for only one programmable processor in a system. In this view, functions that require hardware acceleration—such as 3D rendering—are best performed by fixed-function accelerators. In time, as the PC's central processor becomes faster, less opportunity will remain for media processors. In the near term, though, there appears to be a clear opportunity for such processors to boost PC capabilities for a modest incremental cost.

Embedded processors enable digital consumer electronics

Embedded microprocessors rarely bask in public attention or earn huge profits, but manufacturers produce them in enormous volume and in great diversity. Because software compatibility is not as driving a force as in the desktop market, the embedded market allows more architectures to survive.

Early embedded microprocessor applications were control oriented: Traffic-light and elevator controllers are the classic examples. As microprocessor performance increased, the range of tasks that processors can handle broadened. The vast majority of embedded applications don't demand any more performance than low-cost 8-bit—or even 4-bit—processors offer. Figure 7 shows that, as a result, the bulk of the volume remains with these older devices, which continue to evolve by adding more on-chip peripherals and memory. Ancient 4-bit processors have remained surprisingly popular, but new designs rarely use them because low-end 8-bit devices have dropped

to very low prices and are easier to program. Even so, 4-bit chips—long considered obsolete by most observers—are only now beginning to fade away and will continue shipping more than a billion units per year through the end of the decade.

Some automotive engine controllers, as well as disk and network cards for PCs, use 16-bit embedded processors. (Note that Figure 7 defines 16-bit processors by their external bus width, so it includes in this category many chips with 32-bit internal designs.) Many of these applications are moving to the 32-bit level as application demands increase and 32-bit processor prices drop. In the long run, 8-bit embedded processors will continue to serve the most cost-sensitive applications, while most others gravitate toward 32-bit processors, leaving little room for other sizes.

The most exciting application area for embedded processors is digital consumer electronics. Digital control came to hi-fi equipment years ago, replacing knobs and dials with push buttons and displays. Video cassette recorders gave the microprocessor more sophisticated control functions, but poor user interfaces left most users unwilling to invest the time to learn the new functions. Digital answering machines and compact disc players put the microprocessor in the signal path, marking the beginning of the end of tape for audio storage. Later this year, DVD technology will move video into the digital domain as well.

Video games—which are actually limited-function computers—are the highest volume non-PC applications for 32-bit processors today. Sega's success with its Saturn and Genesis video games, which use Hitachi's SH series processors, has catapulted this relatively recent RISC processor to the third highest volume 32-bit architecture, behind only the venerable x86 and 68000. Sony's PlayStation uses a 32-bit Mips processor, and one video game—the Nintendo 64—uses a 64-bit Mips processor.

Electronic organizers and personal digital assistants (PDAs) are promising application categories for 32-bit embedded processors. Figure 8 (next page) shows one example of a processor designed for such applications. Today's organizers are truly embedded applications; little or no third-party software is available for them. PDAs such as Apple's Newton, on the other hand, are new computing platforms and do depend on third-party software. So far, the success of organizers has been limited to low-cost, limited-function devices, while more-capable PDAs have been successful only in vertical markets. As the technology develops, however, handheld computing devices could become an even bigger industry (in units) than personal computers. Microsoft's new Windows CE will give this application category a big boost.

Internet opens new opportunities

The Internet—in particular the World Wide Web—is creating new classes of consumer computing devices. In fact, it could make PDAs far more compelling, once devices with larger, more readable screens are available. The Web makes vast amounts of information available, significantly increasing the value of a computing device to the average consumer. Early PDA makers hoped to build their own networks and services to offer information such as city guides and restaurant reviews; the Web will do a far better job of providing this

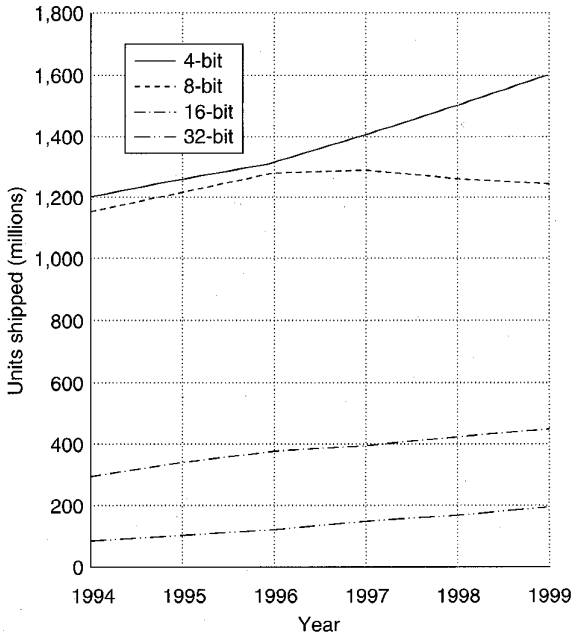


Figure 7. Embedded microprocessor and microcontroller shipments by word size. (*Source: The Information Architects, Mountain View, Calif., 1996*)

content with virtually no investment by the device makers.

While a large-screen PDA might make a great Web access device, it won't be cost-effective until there are major price/performance advances in color flat-panel displays. In the near term, many companies are building Web terminals that connect to televisions for display. WebTV is one company leading the pack in this arena. Its device, built by licensees, uses a Mips R4000-derivative processor created by design-house-turned-fabless-chip-maker Quantum Effect Design (Santa Clara, Calif.). The Web is becoming a central information resource that could eliminate printed phone books and newspaper classified ads; provide customer service and order-processing links to businesses of all kinds; and eventually become a primary delivery mechanism for news and entertainment. As this happens, Web access devices could become a major new class of embedded microprocessor applications.

Advocates of PCs have reacted to Web terminals, not surprisingly, with scorn and derision. After all, they represent a potentially major threat to future PC market growth. If the Web achieves its potential, however, the reality is that easy-to-use, minimum-cost devices that focus on Web access will be successful.

This trend is significant for microprocessor makers, because it breaks the application-software stranglehold the x86 architecture has had on the PC market. A Web browser can run on nearly any architecture. Even applications loaded over the Web can be processor independent if they are written in Java. The Internet, which has already created major markets for 32-bit embedded processors in routers and other network

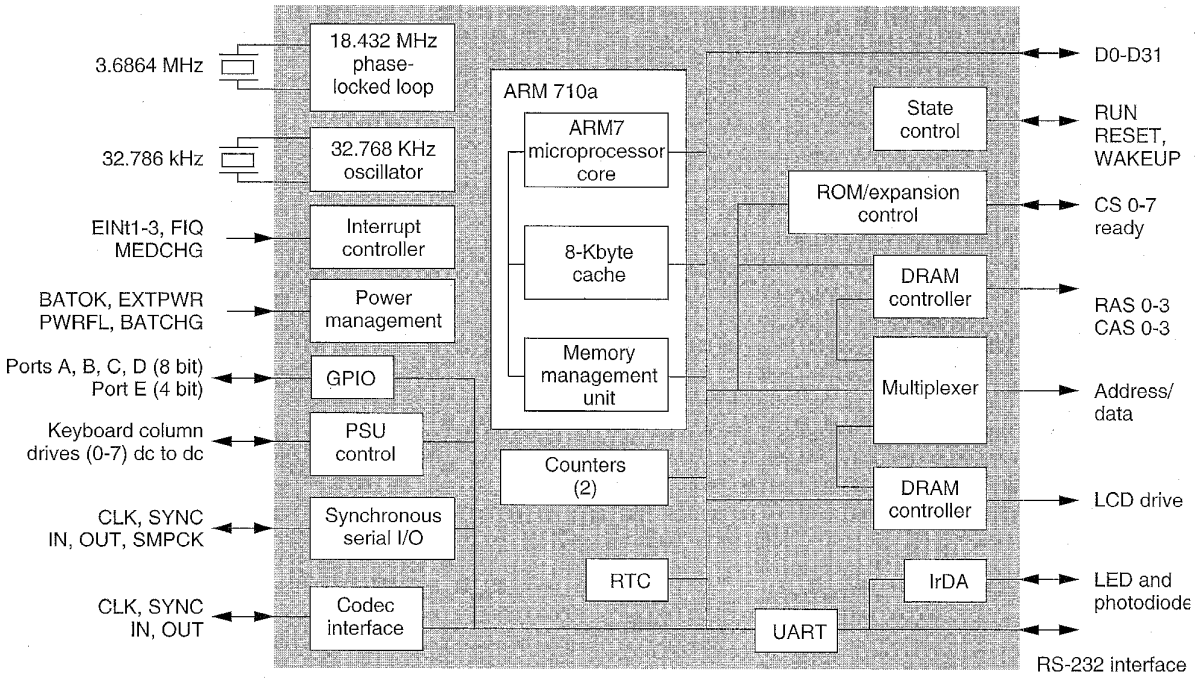


Figure 8. Block diagram of Cirrus Logic's CL-PS7110 integrated processor with ARM7 core.

infrastructure elements, could be a significant enabler of broader competition in the microprocessor business.

Embedded processors proliferate

Table 4 summarizes the key features of a few of the more than one hundred 32- and 64-bit embedded processors now available. As application demands and the competitive environment have changed, architectures have evolved. Digital's StrongARM is a stunningly fast derivative of the power-miserly but not especially fast ARM architecture. Hitachi's new SuperH series has a wide range of devices, of which the table lists only one. Similarly, Motorola and IBM are each producing numerous PowerPC variations for embedded control applications.

Motorola is the champion of embedded processor proliferations, with uncounted 68000 variations. Now it has even modified the base instruction set architecture to produce the RISC-like ColdFire subset. NEC, along with IDT and LSI Logic, is pushing the Mips architecture into embedded applications; Table 4 shows only one of many options. Intel continues to develop its 960 series, which is successful in some markets but shows little sign of progress in the expanding market for low-cost 32-bit processors. (The PC market is a formidable distraction for Intel.)

As high-performance embedded processors move into consumer electronics, low power consumption becomes as important as low price. In portable applications, the value of low power is obvious: longer battery life or smaller, lighter batteries. Even in nonportable consumer applications, however, low power consumption is important, because it reduces the cost of power supplies and eliminates the need for a fan.

These changes in the embedded market have led to major shifts in market share. As Figure 1 shows, Hitachi's SH series has come from nowhere to lead 32-bit RISC processor shipments on the strength of Sega's video games and other consumer applications. Meanwhile, Intel's more traditional 960 series, once the industry leader, has stagnated. AMD has entirely stopped future development of its 29000 family, once the 960's top competitor.

Customization for embedded applications

As transistor counts in chips selling for under $100 (and eventually under $30) skyrocket to millions—and soon to tens of millions—processors for PCs will continue to use most of these transistors to increase performance. For most embedded applications, however, the demands for ever-higher performance just aren't there. Instead, embedded-application designers would like to reduce system costs by integrating more functions on the same chip with the microprocessor. The logical end point of this evolution is a complete system on a chip. Technology is reaching a point where chips can integrate even significant amounts of memory. For example, eliminating half the DRAM array from a 64-Mbit DRAM still leaves 4 Mbytes of memory and room for millions of logic transistors.

As embedded microprocessors evolve toward systems on a chip, they inevitably become more specialized. Different applications have different needs for memory, peripheral controllers, and interfaces to the external world. The desire for

Table 4. Key features of selected embedded microprocessors. *(Source: Vendors except where noted)*

Feature	Digital SA-110	VLSI ARM710	NEC R4300	Hitachi SH7604	IBM PPC 403GA	Motorola 860DC	Motorola 68EC040	Motorola CF5102	Intel 960JA	Intel 960HT
Architecture	ARM	StrongARM	Mips	SuperH	PowerPC	PowerPC	68000	ColdFire	i960	i960
Clock rate (MHz)	200	40	133	20	33	40	40	25	33	60
Instr./data cache size (Kbytes)	16/16	8/8	16/8	4/4	2/1	4/4	4/4	2/1	2/1	16/8
FPU	No	No	Yes	No	No	No	Yes	No	No	No
MMU	Yes	Yes	Yes	No	No	Yes	Yes	No	No	No
Bus frequency (MHz)	66	40	66	20	33	40	40	25	33	20
MIPS†	230	36	160*	20	41	52	44	27	28	100*
Voltage	2.0/3.3**	5	3.3	3.3	3.3	3.3	5	3.3	3.3	3.3
Power (typical, mW)	900	424	2,200	200	265	900	4,500	900	500	4,500
MIPS/watt	239	85	73	100	155	58	10	30	56	22
MIPS/price	4.30	1.04	5.00	0.24	1.05	0.51	0.59	N.A.	0.76	0.79
Transistors (millions)	2.1	0.6	1.7	0.45	0.58	1.8	1.2	N.A.	0.75	2.3
Process (µm/layers)	0.35/3	0.6/2	0.35/3	0.8/2	0.5/3	0.5/3	0.65/3	0.6/3	0.8/3	0.6/4
Die size (mm^2)	50	34	45	82	39	25	163	N.A.	64	100
Estimated mfg. cost*	$18	$9	$11	$7	$14	$20	$30	$9	$8	$34
Availability	Now	Now	Now	Now	Now	Now	Now	Now	Now	Now
List price (10,000s)	$49	$28	$32	$27	$28	$102	$75	$25	$37	$126

* MicroDesign Resources estimate ** Core/bus voltage † MIPS rating as supplied by vendor, based on Dhrystone 2.1

highly integrated system chips is increasing the demand for building-block microprocessors that can function as parts of application-specific integrated circuits (ASICs). Many of the leading microprocessor vendors are not major ASIC suppliers, however, nor are they set up to customize chips for every customer. Indeed, eliminating the need to do so was a key benefit of the microprocessor in the first place.

LSI Logic is one company that has pioneered the design of ASICs with microprocessor cores. Many other companies, including Texas Instruments, IBM Microelectronics, VLSI Technology, and NEC, are also aggressively developing this technology. Not only must these companies have a range of microprocessor cores available, but they must provide a variety of other complex building blocks, such as MPEG decoders and graphics engines, as well as the software tools to design, debug, verify, and test the chips. In the future embedded-processor market, these factors may be more important than the processor cores themselves.

In this world of core-based ASICs, some microprocessor cores are becoming near commodities. Advanced RISC Machines (ARM) in the UK has licensed its core designs widely, and many companies offer ARM cores as part of their ASIC libraries. Mips has also licensed its cores widely, though not as widely as ARM, and Sparc cores have a few licensees.

Table 5. Originators and licensees of RISC processors for embedded applications.

Architecture	Originator	Licensees
ARM	ARM Ltd.	Asahi Kasei Microsystems (AKM), Alcatel, Atmel, Cirrus Logic, Digital, GEC Plessey, LG Semicon, NEC, Oki, Samsung, Sharp, Symbios Logic, Texas Instruments, VLSI Technology, Yamaha
ColdFire	Motorola	Mitsubishi
Mips	Mips Technologies	Integrated Device Technology, LSI Logic, NEC, NKK, Philips, QED, Sony, Toshiba
PowerPC	IBM Microelectronics	Mitsubishi
SuperH	Hitachi	VLSI Technology
Sparc	Sun Microelectronics	C-Cube, Fujitsu, Hyundai, Matra MHS, Scientific Atlanta, TGI

Table 5 shows CPU architectures that companies have licensed to chip and equipment makers for embedded applications. Motorola continues to keep most of its cores proprietary and is gradually allowing more and more customer involvement in the design process.

Packaging is another key area that needs improvement. As designers put more functions on a chip, the chips need more input/output pins. Today's common plastic quad flat packs offer a cost per pin around 2 cents, but can't provide pin counts much beyond 200. High-pin-count pin grid arrays typically have costs around 10 cents per pin—leading to a $50 package for a 500-pin device. New packaging technologies, such as plastic ball grid arrays and various chip-scale packages, promise high-pin-count packages with costs approaching a penny per

Web resources on microprocessors

Independent sites

Hitex Software's numerical chip directory: www.hitex.com/chipdir
IEEE Micro: www.computer.org/pubs/micro/micro.htm
John Bayko's "Great Microprocessors of the Past and Present": www.cs.uregina.ca:80/~bayko/cpu.html
Ken Polsson's Timeline of Microcomputers: www.islandnet.com/~kpolsson/comphist.htm
MicroDesign Resources: www.chipanalyst.com
SPEC benchmark results: www.specbench.org
The Actual Size Processor Page: www.gulf.net/~stone/processor
UC Berkeley CPU Info Center: infopad.eecs.berkeley.edu/CIC

Microprocessor manufacturers

AMD: www.amd.com
ARM: www.arm.com
Cyrix: www.cyrix.com
Chromatic Research: www.chromatic.com
Digital: www.digital.com
Fujitsu: www.fujitsu.com
Hewlett-Packard (PA-RISC): www.hp.com/wsg/strategies/strategy.html
Hitachi: www.hitachi.com
IBM Microelectronics: www.chips.ibm.com
Integrated Device Technology: www.idt.com
Intel: www.intel.com
LSI Logic: www.lsilogic.com
Microchip: www.microchip.com
Mips: www.mips.com
Motorola: www.mot.com/SPS/General
National Semiconductor: www.nsc.com
NEC: www.nec.com
Oki: www.oki.com
Philips: www.semiconductors.philips.com
QED: www.qedinc.com
Ross Technology: www.ross.com
Siemens Semiconductor: www.scf.siemens.com
Texas Instruments: www.ti.com
VLSI Technology: www.vlsi.com
Zilog: www.zilog.com

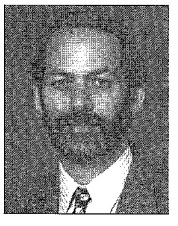
pin in the next few years. If this comes to pass, it would be a significant enabling technology for highly integrated, low-cost chips. Sometimes the silicon seems like the easy part!

WHERE NEXT? After 25 years of development, advancements in microprocessor technology show no signs of slowing down. The pace of new architecture introductions has slowed, especially in the desktop market, but new implementations are coming out at record rates. Rapidly increasing transistor counts and clock speeds are challenging designers to innovate continually to deliver the most value from the technology. And as the Web has so vividly demonstrated, major new applications may be just around the corner—but are extraordinarily difficult to forecast.

PCs are becoming potent communication and entertainment devices and are moving into homes in a big way. At the same time, many new consumer electronics devices—from Web terminals to DVD players—are becoming available. It is hard to predict just which devices will succeed. But it is a sure bet that ever-advancing microprocessor technology will be crucial to the products enabling the much-discussed convergence of computing, communication, and entertainment. This ensures the microprocessor's role at the heart of the electronics industry for the next 25 years or longer.

Acknowledgments

This article would not have been possible without the combined efforts of the MicroDesign Resources analyst team. In particular, I'd like to thank Linley Gwennap, who is both our expert on high-performance microprocessor design and the leader of the team; Jim Turley, who tracks embedded microprocessors and their applications; and Yong Yao, Peter Glaskowsky, and Steve Hammond, who track a range of PC hardware technologies.



Michael Slater is the founder and editorial director of MicroDesign Resources and publisher of *Microprocessor Report,* an industry newsletter. He also organizes the Microprocessor Forum and PC Tech Forum conferences and consults for leading microprocessor and computer companies on new product strategies. A well-known speaker on microprocessor technology and system trends, Slater is also a columnist for *Electronic Engineering Times, Nikkei Electronics Asia,* and *Computer Shopper.*

Direct questions concerning this article to Michael Slater, MicroDesign Resources, 874 Gravenstein Hwy. S., Suite 14, Sebastopol, CA 95472; mslater@mdr.zd.com.

Reader Interest Survey

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

Low 160 Medium 161 High 162