Chapter 6

Microprocessor Types and Specifications

The brain of the PC is the processor, or Central Processing Unit (CPU). The CPU performs the system's calculating and processing (except for special math-intensive processing in systems that have a math coprocessing unit chip). The processor is easily the most expensive chip in the system. All the IBM PC and PS/2 units and IBM-compatibles use processors that are compatible with the Intel family of chips, although the processors themselves may have been manufactured or designed by various companies, including IBM, Cyrix, and AMD.

The following sections cover the processor chips that have been used in personal computers since the first PC was introduced more than a decade ago. These sections provide a great deal of technical detail about these chips and explain why one type of CPU chip can do more work than another in a given period of time. First, however, you learn about two important components of the processor: the data bus and the address bus.

Processor Specifications

Confusing specifications often are quoted in discussions of processors. The following sections discuss some of these specifications, including the data bus, address bus, and speed. The next section includes a table that lists the specifications of virtually all PC processors.

Data Bus

One of the most common ways to describe a processor is by the size of the processor's data bus and address bus. A *bus* is simply a series of connections that carry common signals. Imagine running a pair of wires from one end of a building to another. If you connect a 110-volt AC power generator to the two wires at any point and place outlets at convenient locations along the wires, you have constructed a power bus. No matter which outlet you plug the wires in to, you have access to the same signal, which in this example is 110-volt AC power.

Any transmission medium that has more than one outlet at each end can be called a bus. A typical computer system has several buses, and a typical processor has two important buses for carrying data and memory-addressing information: the data bus and the address bus.

The processor bus discussed most often is the *data bus*: the bundle of wires (or pins) used to send and receive data. The more signals that can be sent at the same time, the more data can be transmitted in a specified interval and, therefore, the faster the bus.

Data in a computer is sent as digital information, consisting of a time interval in which a single wire carries 5 volts to signal a 1 data bit or 0 volts to signal a 0 data bit. The more wires you have, the more individual bits you can send in the same time interval. A chip such as the 286, which has 16 wires for transmitting and receiving such data, has a 16-bit data bus. A 32-bit chip such as the 486 has twice as many wires dedicated to simultaneous data transmission as a 16-bit chip and can send twice as much information in the same time interval as a 16-bit chip.

A good way to understand this flow of information is to consider a highway and the traffic it carries. If a highway has only one lane for each direction of travel, only one car at a time can move in a certain direction. If you want to increase traffic flow, you can

Table 6.1 Intel Processor Specifications

Processor	CPU Clock	Std. Voltage	Internal Register Size	Data Bus Width	Address Bus Width	Maximum Memory
8088	1x	5v	16-bit	8-bit	20-bit	1M
8086	1x	5v	16-bit	16-bit	20-bit	1M
286	1x	5v	16-bit	16-bit	24-bit	16M
386SX	1x	5v	32-bit	16-bit	24-bit	16M
386SL	1x	3.3v	32-bit	16-bit	24-bit	16M
386DX	1x	5v	32-bit	32-bit	32-bit	4G
486SX	1x	5v	32-bit	32-bit	32-bit	4G
486SX2	2x	5v	32-bit	32-bit	32-bit	4G
487SX	1x	5v	32-bit	32-bit	32-bit	4G
486DX	1x	5v	32-bit	32-bit	32-bit	4G
486SL**	1x	3.3v	32-bit	32-bit	32-bit	4G
486DX2	2x	5v	32-bit	32-bit	32-bit	4G
486DX4	2-3x	3.3v	32-bit	32-bit	32-bit	4G
Pentium 60/66	1x	5v	32-bit	64-bit	32-bit	4G
Pentium 90/100	1.5-2x	3.3v	32-bit	64-bit	32-bit	4G

**The 386SL contains an integral-cache controller, but the cache memory must be provided outside the chip.*
***The 486SL processor has been discontinued; instead, Intel now markets SL Enhanced versions of the SX, DX, and DX2 processors. These pr*
FPU = Floating-Point Unit (math coprocessor)
WT = Write-Through cache (caches reads only)
WB = Write-Back cache (caches both reads and writes)

add another lane so that twice as many cars pass in a specified time. You can think of an 8-bit chip as being a single-lane highway because with this chip, one byte flows through at a time. (One byte equals eight individual bits.) The 16-bit chip, with two bytes flowing at a time, resembles a two-lane highway. To move a large number of automobiles, you may have four lanes in each direction. This structure corresponds to a 32-bit data bus, which has the capability to move four bytes of information at a time.

Just as you can describe a highway by its lane width, you can describe a chip by the width of its data bus. When you read an advertisement that describes a computer system as being a 16-bit or 32-bit system, the ad usually is referring to the data bus of the CPU. This number provides a rough idea of the performance potential of the chip (and, therefore, the system).

Table 6.1 lists the specifications, including the data-bus sizes, for the Intel family of processors used in IBM and compatible PCs.

Internal Registers

The size of the internal register is a good indication of how much information the processor can operate on at one time. Most advanced processors today—all chips from the 386 to the Pentium—use 32-bit internal registers.

Integral Cache	Cache Type	Burst Mode	Integral FPU	No. of Transistors	Date Introduced
No	—	No	No	29,000	June 1979
No	—	No	No	29,000	June 1978
No	—	No	No	134,000	Feb. 1982
No	—	No	No	275,000	June 1988
0K*	WT	No	No	855,000	Oct. 1990
No	—	No	No	275,000	Oct. 1985
8K	WT	Yes	No	1,185,000	April 1991
8K	WT	Yes	No	1,185,000	April 1994
8K	WT	Yes	Yes	1,200,000	April 1991
8K	WT	Yes	Yes	1,200,000	April 1989
8K	WT	Yes	Optional	1,400,000	Nov. 1992
8K	WT	Yes	Yes	1,100,000	March 1992
16K	WT	Yes	Yes	1,600,000	Feb. 1994
2×8K	WB	Yes	Yes	3,100,000	March 1993
2×8K	WB	Yes	Yes	3,300,000	March 1994

X2 processors. These processors are available in both 5v and 3.3v versions and include power-management capabilities.

Some processors have an internal data bus (made up of data paths and of storage units called *registers*) that is different from the external data bus. The 8088 and 386SX are examples of this structure. Each chip has an internal data bus that is twice the width of the external bus. These designs, which sometimes are called *hybrid designs*, usually are low-cost versions of a "pure" chip. The 386SX, for example, can pass data around internally with a full 32-bit register size; for communications with the outside world, however, the chip is restricted to a 16-bit-wide data path. This design enables a systems designer to build a lower-cost motherboard with a 16-bit bus design and still maintain compatibility with the full 32-bit 386.

Internal registers often are larger than the data bus, which means that the chip requires two cycles to fill a register before the register can be operated on. For example, both the 386SX and 386DX have internal 32-bit registers, but the 386SX has to "inhale" twice (figuratively) to fill them, whereas the 386DX can do the job in one "breath." The same thing would happen when the data is passed from the registers back out to the system bus.

The Pentium is an example of the opposite situation. This chip has a 64-bit data bus but only 32-bit registers—a structure that may seem to be a problem until you understand that the Pentium has two internal 32-bit pipelines for processing information. In many ways, the Pentium is like two 32-bit chips in one. The 64-bit data bus provides for very efficient filling of these multiple registers.

Address Bus

The *address bus* is the set of wires that carry the addressing information used to describe the memory location to which the data is being sent or from which the data is being retrieved. As with the data bus, each wire in an address bus carries a single bit of information. This single bit is a single digit in the address. The more wires (digits) used in calculating these addresses, the greater the total number of address locations. The size (or width) of the address bus indicates the maximum amount of RAM that a chip can address.

The highway analogy can be used to show how the address bus fits in. If the data bus is the highway, and if the size of the data bus is equivalent to the number of lanes, the address bus relates to the house number or street address. The size of the address bus is equivalent to the number of digits in the house address number. For example, if you live on a street in which the address is limited to a two-digit (base 10) number, no more than 100 distinct addresses (00 to 99) can exist for that street (10 to the power of 2). Add another digit, and the number of available addresses increases to 1,000 (000 to 999) or 10 to the 3rd power.

Computers use the binary (base 2) numbering system, so a two-digit number provides only four unique addresses (00, 01, 10, and 11) calculated as 2 to the power of 2, and a three-digit number provides only eight addresses (000 to 111) which is 2 to the 3rd power. For example the 8086 and 8088 processors use a 20-bit address bus which calculates as a maximum of 2 to the 20th power or 1,048,576 bytes (1M) of address locations. Table 6.2 describes the memory-addressing capabilities of Intel processors.

Table 6.2 Intel Processor Memory-Addressing Capabilities

Processor Family	Address Bus	Bytes	Kilobytes	Megabytes	Gigabytes
8088/8086	20 bits	1,048,576	1,024	1	—
286, 386SX	24 bits	16,777,216	16,384	16	—
386DX, 486, Pentium	32 bits	4,294,967,296	4,194,304	4,096	4

The data bus and address bus are independent, and chip designers can use whatever size they want for each. Usually, however, chips with larger data buses have larger address buses. The sizes of the buses can provide important information about a chip's relative power, measured in two important ways. The size of the data bus is an indication of the information-moving capability of the chip, and the size of the address bus tells you how much memory the chip can handle.

Processor Speed Ratings

A common misunderstanding about processors is their different speed ratings. This section covers processor speed in general and then provides more specific information about Intel processors.

A computer system's *clock speed* is measured as a frequency, usually expressed as a number of cycles per second. A crystal oscillator controls clock speeds, using a sliver of quartz in a small tin container. As voltage is applied to the quartz, it begins to vibrate (oscillate) at a harmonic rate dictated by the shape and size of the crystal (sliver). The oscillations emanate from the crystal in the form of a current that alternates at the harmonic rate of the crystal. This alternating current is the *clock signal*. A typical computer system runs millions of these cycles per second, so speed is measured in megahertz (MHz). (One hertz is equal to one cycle per second.)

Note

The hertz was named for the German physicist Heinrich Rudolph Hertz. In 1885, Hertz confirmed the electromagnetic theory through experimentation which states that light is a form of electromagnetic radiation and is propagated as waves.

A single cycle is the smallest element of time for the processor. Every action requires at least one cycle and usually multiple cycles. To transfer data to and from memory, for example, an 8086 chip needs four cycles plus wait states. (A *wait state* is a clock tick in which nothing happens to ensure that the processor isn't getting ahead of the rest of the computer.) A 286 needs only two cycles plus any wait states for the same transfer.

The time required to execute instructions also varies. The original 8086 and 8088 processors take an average of 12 cycles to execute a single instruction. The 286 and 386 processors improve this rate to about 4.5 cycles per instruction; the 486 drops the rate further, to 2 cycles per instruction. The Pentium includes twin instruction pipelines and other improvements that provide for operation at 1 cycle per average instruction.

Different instruction execution times (in cycles) makes it difficult to compare systems based purely on clock speed, or number of cycles per second. One reason why the 486 is so fast is that it has an average instruction-execution time of 2 clock cycles. Therefore, a 100 MHz Pentium is about equal to a 200 MHz 486, which is about equal to a 400 MHz 386 or 286, which is about equal to a 1,000 MHz 8088. As you can see, you have to be careful in comparing systems based on pure MHz alone; many other factors affect system performance.

How can two processors that run at the same clock rate perform differently, with one running "faster" than the other? The answer is simple: efficiency. Suppose that you are comparing two engines. An engine has a crankshaft revolution, called a cycle. This cycling time is measured in revolutions per minute (RPM). If two engines run the same maximum RPM, they should run the car at the same speed, right?

Wrong! Suppose that you're shopping for a fast family car, and you decide to compare the Ford SHO Taurus against the Chevrolet Impala SS. You stop first at the Ford dealership and look at the '95 SHO. You ask the dealer, "What's the engine speed at maximum power output?" The dealer tells you that the SHO has a 3.2-liter, 6-cylinder engine that puts out 220 HP at 6,200 RPM. You are impressed by the high RPM rating, and you record the information.

Then you see the Chevrolet dealer, who steers you toward a '95 Impala SS. You ask the same question about engine power and speed, and the dealer tells you that the Corvette-style LT1 5.7-liter, V-8 engine used in the SS puts out 260 HP at 5,000 RPM. You now figure that because the Ford engine turns 1,200 RPM higher when producing maximum power, it will propel that car much faster than the Chevrolet, the engine of which is turning only 5,000 RPM.

Actually, the car with the higher power output is faster, assuming that the cars weigh the same (which they do not). As you can see, these types of specification comparisons can be difficult to manage because of all of the other variables that can enter into the comparison. I would not compare two computer systems based solely on MHz any more than I would compare two cars on the basis of engine RPM.

You can see that comparing the performance of two vehicles based solely on engine RPM is inaccurate. You never would make such a comparison, because you know that many more factors than just engine speed determine vehicle speed and acceleration capability.

Unfortunately, we often make the same type of poor comparison in evaluating computers. Using engine RPM to compare how fast two cars can run is similar to using MHz to compare how fast two computers can run. A better specification to use when you are comparing the two vehicles would be engine horsepower, which is a measurement of the amount of work that each engine can perform. Then you would have to adjust the horsepower figure for the weight of the vehicle, the coefficient of drag, drive-line gearing, parasitic losses, and so on. In effect, too many other variables are involved for you to make any simplistic comparison, even if you first picked a more meaningful specification to compare than engine redline. The best way to evaluate which of the two vehicles is faster is through a road test. In a computer, the equivalent is taking some of your software and running *benchmarks*, or comparative performance tests.

The big V-8 engine in the Chevrolet does more work in each crankshaft revolution (or cycle) than the 6-cylinder engine in the Ford. In the same manner, a 286 or 386 can perform much more work in a single CPU cycle than an 8088 can; it's simply more efficient. As you can see, you must be careful in comparing MHz to MHz, because much more is involved in total system performance.

Comparing automobile engines by virtue of their horsepower output is a much more valid comparison than just RPM. What is needed is a sort of horsepower measurement for processors. To compare processors more accurately based on comparative "horsepower," Intel has devised a specific series of benchmarks that can be run against Intel chips to produce a relative gauge of performance. This is called the ICOMP (Intel COmparative Microprocessor Performance) index. The following tables show the relative power, or ICOMP index, for several processors.

Table 6.3 Intel ICOMP Index Ratings

Processor	ICOMP Index
386SX-16	22
386SX-20	32
386SX-25	39
386SL-25	41
386DX-25	49
486SX-16	63
386DX-33	68
486SX-20	78
486SX-25	100
486DX-25	122
486SX-33	136
486DX-33	166
486SX2-50	180
486DX2-50	231
486DX-50	249
486DX2-66	297
486DX4-75	319
486DX4-100	435
Pentium 60	510
Pentium 55	567
Pentium 90	735
Pentium 100	815

The ICOMP index is derived from several independent benchmarks and is a stable indication of relative processor performance. Floating-point calculations are weighed in the ICOMP rating, so processors that have a built-in FPU (Floating-Point Unit) always have some advantage over those that do not.

Another factor in CPU performance is clock speed. Clock speed is a function of a system's design and usually is controlled by an oscillator, which in turn is controlled by a quartz crystal. Typically, you divide the crystal-oscillation frequency by some amount to obtain the processor frequency. The divisor amount is determined by the original design of the processor (Intel), by related support chips, and by how the motherboard was designed to use these chips together as a system. In IBM PC and XT systems, for example, the main crystal frequency is 14.31818 MHz, which is divided by 3 by an 8284 clock generator chip to obtain a 4.77 MHz processor clock speed. In an IBM AT system, the crystal speeds are 12 or 16 MHz, which is divided by 2 internally inside the 80286 to produce a 6 or 8 MHz processor clock speed, respectively.

If all other variables are equal—including the type of processor, the number of wait states (empty cycles) added to memory accesses, and the width of the data bus—you can compare two systems by their respective clock rates. Be careful with this type of comparison, however; certain variables (such as the number of wait states) can greatly influence the speed of a system, causing the unit with the lower clock rate to run faster than you expect, or causing the system with a numerically higher clock rate to run slower than you think it should. The construction and design of the memory subsystem can have an enormous effect on a system's final execution speed.

In building a processor, a manufacturer tests it at different speeds, temperatures, and pressures. After the processor is tested, it receives a stamp indicating the maximum safe speed at which the unit will operate under the wide variation of temperatures and pressures encountered in normal operation. The rating system usually is simple. For example, the top of the processor in my system is marked like this:

    A80486DX2-66

The *A* is Intel's indicator that this chip has a Ceramic Pin Grid Array form factor, which describes the physical packaging of the chip. The *80486DX2* is the part number, which identifies this processor as a clock-doubled 486DX processor. The *-66* at the end indicates that this chip is rated to run at a maximum speed of 66 MHz. Because of the clock doubling, the maximum motherboard speed is 33 MHz. This chip would be acceptable for any application in which the chip runs at 66 MHz or slower. For example, you could use this processor in a system with a 25 MHz motherboard, in which case the processor would happily run at 50 MHz.

Sometimes, however, the markings don't seem to indicate the speed. In the 8086, for example, *-3* translates to 6 MHz operation. This marking is more common in some of the older chips, manufactured before some of the marking standards used today were standardized.

A manufacturer sometimes places the CPU under a heat sink, which prevents you from reading the rating printed on the chip. (A *heat sink* is a metal device that draws heat away from an electronic device.) Although having a heat sink generally is a good idea, Intel designs its chips to run at rated speed without a heat sink.

Intel Processors

IBM-compatible computers use processors manufactured primarily by Intel. Some other companies, such as Cyrix and AMD, have reverse-engineered the Intel processors and have begun making their own compatible versions. IBM also manufactures processors for some of its own systems as well as for installation in boards and modules sold to others. The IBM processors are not reverse-engineered, but are produced in cooperation with and under license from Intel. The terms of IBM's agreement also allows IBM to make modifications and improvements to the basic Intel design. IBM uses these processors in their own systems, but also is allowed to sell the processors in boards they manufacture for other companies. IBM is not allowed to sell these chips raw, but must always install them in some type of board assembly.

Knowing the processors used in a system can be very helpful in understanding the capabilities of the system, as well as in servicing it. To fully understand the capabilities of a system and perform any type of servicing, you must at least know the type of processor that the system uses.

8088 and 8086 Processors

The original IBM PC used an Intel CPU chip called the 8088. The original 8088 CPU chip ran at 4.77 MHz, which means that the computer's circuitry drove the CPU at a rate of 4,770,000 *ticks*, or computer heartbeats, per second. Each tick represents a small amount of work—the CPU executing an instruction or part of an instruction—rather than a period of elapsed time.

In fact, both the 8088 and 8086 take an average 12 cycles to execute the average instruction. The 8088 has an external data bus 8 bits wide, which means that it can move 8 bits (individual pieces) of information into memory at a time. The 8088 is referred to as a 16-bit processor, however, because it features internal 16-bit-wide registers and data paths. The 8088 also has a 20-bit address bus, which enables the system to access 1M of RAM. Using the 8088, a manufacturer could build a system that would run 16-bit software and have access to 1M of memory while keeping the cost in line with then-current 8-bit designs. Later, IBM used the 8088 chip in the PC/XT computer.

IBM used the 8088 to put together the original IBM PC 5150-001, which sold for $1,355 with 16K of RAM and no drives. A similarly configured Apple II system, the major competition for the original PC, cost about $1,600.

The 8088 eventually was redesigned to run at 8 MHz—nearly double the speed of the original PC. The speed at which the processor operates has a direct effect on the speed of program execution. Later sections of this chapter cover the speeds of the CPU chips that are successors to the 8088.

Note

The real mode of 286 and higher CPU chips refers to the mode that these advanced chips use to imitate the original 8088 chip in the first PC. Real mode is used by 286 and higher CPU chips to run a single DOS program at a time, just as though systems based on these powerful chips are merely faster PCs. The additional modes of 286 and higher CPU chips are covered in subsequent sections of this chapter.

Computer users sometimes wonder why a 640K conventional-memory barrier exists if the 8088 chip can address 1M of memory. The conventional-memory barrier exists because IBM reserved 384K of the upper portion of the 1,024K (1M) address space of the 8088 for use by adapter cards and system BIOS (a computer program permanently "burned into" the ROM chips in the PC). The lower 640K is the conventional memory in which DOS and software applications execute.

In 1976, before the 8088 chip, Intel made a slightly faster chip named the 8086. The 8086, which was one of the first 16-bit chips on the market, addressed 1M of RAM. The design failed to catch on, however, because both the chip and a motherboard designed for the chip were costly. The cost was high because the system needed a 16-bit data bus rather than the less expensive 8-bit bus. Systems available at that time were 8-bit, and users apparently weren't willing to pay for the extra performance of the full 16-bit design. Therefore, Intel introduced the 8088 in 1978. Both the 8086 and the 8088 CPU chips are quite slow by today's standards.

IBM largely ignored the 8086 CPU chip until it manufactured the first PS/2 Models 25 and 30. Systems produced by many other manufacturers, such as the Compaq Deskpro and the AT&T 6300, had been using the 8086 for some time. The capability of the 8086 to communicate with the rest of the system at 16 bits gives it about a 20 percent increase in processing speed over an 8088 with an identical speed (in MHz). This improvement is one reason why IBM can claim that the 8 MHz, 8086-based Model 30 is 2-1/2 times faster than the 4.77 MHz, 8088-based PC or XT, even though 8 MHz is not more than twice the clock speed. This claim is the first indication of what a CPU chip with a wider data path can mean in terms of speed improvements.

80186 and 80188 Processors

After Intel produced the 8086 and 8088 chips, it turned its sights toward producing a more powerful chip with an increased instruction set. The company's first efforts along this line—the 80186 and 80188—were unsuccessful. But incorporating system components into the CPU chip was an important idea for Intel, because it led to faster, better chips, such as the 286.

The relationship between the 80186 and 80188 is the same as that of the 8086 and 8088; one is a slightly more advanced version of the other. Compared CPU to CPU, the 80186 is almost the same as the 8088 and has a full 16-bit design. The 80188 (like the 8088) is a hybrid chip that compromises the 16-bit design with an 8-bit external communications interface. The advantage of the 80186 and 80188 is that they combine on a single chip

15 to 20 of the 8086–8088 series system components, a fact that can greatly reduce the number of components in a computer design. The 80186 and 80188 chips are used for highly intelligent peripheral adapter cards, such as network adapters.

Although the 80186 and 80188 did provide some new instructions and capabilities, not much in those chips was new compared with the improvements that came later, in the 286 and higher chips. The 80186 and 80188 chips were difficult for systems designers to use in manufacturing systems that were compatible with the IBM PC. For example, these chips had DMA (Direct Memory Access) and Interrupt controllers built-in, but they were incompatible with the external controllers required for an IBM-compatible PC design. Slight differences in the instruction sets also caused problems when the 80186 and 80188 were supposed to emulate 8086 and 8088 chips. In addition to compatibility problems,
the chips didn't offer much performance improvement over the earlier 8086 and 8088. In addition, the individual components that the 80186 and 80188 chips were designed to replace had become inexpensive, which made the 80186 and 80188 chips less attractive.

286 Processors

The Intel 80286 (normally abbreviated as 286) processor did not suffer from the compatibility problems that damned the 80186 and 80188. The 286 chip, introduced in 1981, is the CPU behind the IBM AT. You also can find 286 chips in IBM's original PS/2 Models 50 and 60 (later PS/2s contain 386 or 486 chips). Other computer makers manufactured what came to be known as *IBM clones*, many of these manufacturers calling their systems AT or AT-class computers.

When IBM developed the AT, it selected the 286 as the basis for the new system because the chip provided much compatibility with the 8088 used in the PC and the XT, which means that software written for those chips should run on the 286. The 286 chip is many times faster than the 8088 used in the XT, and it offered a major performance boost to PCs used in businesses. The processing speed, or *throughput*, of the original AT (which ran at 6 MHz) was five times greater than that of the PC running at 4.77 MHz.

For several reasons, 286 systems are faster than their predecessors. The main reason is that 286 processors are much more efficient in executing instructions. An average instruction takes 12 clock cycles on the 8086 or 8088 but an average 4.5 cycles on the 286 processor. Additionally, the 286 chip can handle up to 16 bits of data at a time through an external data bus that is twice the size of the 8088.

Another reason why personal computing received a major boost from the 286 chip is clock speed. AT-type systems are based on 6, 8, 10, 12, 16, and 20 MHz versions of the 286 chip. Earlier processors typically are available in versions only up to 8 MHz. Even if the clock speeds are the same—as in comparing a system that has an 8 MHz 8088 with a system
that has an 8 MHz 286—the 286-based system operates roughly three times faster.

The 286 chip has two modes of operation: real mode and protected mode. The two modes are distinct enough to make the 286 resemble two chips in one. In real mode, a

286 acts essentially the same as an 8086 chip and is fully object-code-compatible with the 8086 and 8088. (A processor with *object-code-compatibility* can run programs written for another processor without modification and execute every system instruction in the same manner.)

In the protected mode of operation, the 286 truly was something new. In this mode, a program designed to take advantage of the chip's capabilities believes that it has access to 1G of memory (including virtual memory). The 286 chip, however, can address only 16M of hardware memory. When a program calls for more memory than physically exists in the system, the CPU swaps to disk some of the currently running code and enables the program to use the newly freed physical RAM. The program does not know about this swapping and instead acts as though 1G of actual memory exists. V*irtual memory* is controlled by the operating system and the chip hardware.

A significant failing of the 286 chip is that it cannot switch from protected mode to real mode without a hardware reset (a warm reboot) of the system. (It can, however, switch from real mode to protected mode without a reset.) A major improvement of the 386 over the 286 is the fact that software can switch the 386 from real mode to protected mode, and vice versa.

When the 286 chip was introduced, Intel said that real mode was created so that much of the 8086- and 8088-based software could run with little or no modification until new software could be written to take advantage of the protected mode of the 286. As with later Intel processors, however, it was a long time before software took advantage of the capabilities of the 286 chip. For example, most 286 systems are used as if they are merely faster PCs. These systems are run in real mode most of the time because the programs were written for DOS, and DOS and DOS programs are limited to real mode. Unfortunately, much of the power of systems based on the 286 chip is unused. In real mode, a 286 chip cannot perform any additional operations or use any extra features designed into the chip.

IBM and Microsoft together began the task of rewriting DOS to run in both real and protected modes. The result was early versions of OS/2, which could run most old DOS programs just as they ran before, in real mode. In protected mode, OS/2 provided true software multitasking and access to the entire 1G of virtual or 16M of physical address space provided by the 286. UNIX and XENIX also were written to take advantage of the 286 chip's protected mode. In terms of mass appeal, however, these operating systems were a limited success.

Little software that took advantage of the 286 chip was sold until Windows 3.0 offered Standard Mode for 286 compatibility, and by that time, the hottest-selling chip was the 386. Still, the 286 was Intel's first attempt to produce a CPU chip that supported *multitasking*, in which multiple programs run at the same time. The 286 is designed so that if one program locks up or fails, the entire system doesn't need a warm boot (reset) or cold boot (power off or on). Theoretically, what happens in one area of memory doesn't affect other programs. Before multitasked programs are "safe" from one another, however, the 286 chip (and subsequent chips) needs an operating system that works cooperatively with the chip to provide such protection.

In a way, this situation leads back to OS/2, which could provide protection but never caught on for the 286 in a big way. Although newer versions of OS/2 offer a graphical user interface similar to that of Windows—and although on 386 and newer systems, OS/2 offers full 32-bit processing for software that is designed to take advantage of it—OS/2 is nowhere near to replacing DOS as the operating system of choice on PCs and is nowhere near as popular as Windows. One reason why is the fact that few OS/2 applications have been developed, compared with the number of DOS and Windows programs.

Protected mode on a 286 enables multiple programs to run at one time only when those programs are specifically written for the operating system (or operating environment). To run several programs at the same time on a 286 in Windows, for example, each active program must be a Windows program (written specifically to run only under Windows).

Because of the virtual-memory scheme of the 286, the size of programs under operating systems such as OS/2 and UNIX can be extremely large. Even though the 286 does not address more than 16M of physical memory, the 286's virtual-memory scheme enables programs to run as though 1G of memory were available. But programs that require a great deal of swapping run slowly, which is why software manufacturers usually indicate the amount of physical RAM needed to run their programs effectively. The more physical memory you install, the faster 286-based systems running OS/2 or UNIX will work.

Windows 3.0, which is not a true operating system because it uses DOS for its underpinnings, provides only poor protection on a 286. Unruly programs still can crash the entire system. Windows 3.1 does a better job of implementing protection on a 286, but it still is far from perfect.

Although UNIX and XENIX provide support for the 286 chip's protected mode, these operating systems have found a following among a small group of extremely high-end computer users, primarily in academic or scientific settings.

386 Processors

The Intel 80386 (normally abbreviated as 386) caused quite a stir in the PC industry because of the vastly improved performance that it brought to the personal computer. Compared with 8088 and 286 systems, the 386 chip offers greater performance in almost all areas of operation.

The 386 is a full 32-bit processor optimized for high-speed operation and multitasking operating systems. Intel introduced the chip in 1985, but the 386 appeared in the first systems in late 1986 and early 1987. The Compaq Deskpro 386 and systems made by several other manufacturers introduced the chip; somewhat later, IBM used the chip in its
PS/2 Model 80. For several years, the 386 chip rose in popularity, which peaked around 1991. Since then, the popularity of the 386 has waned; in the past year or so, it has virtually died out, due to the availability of inexpensive systems based on the 486 and Pentium CPU chips. The 386 had an extended life in the mainstream due in part to the use of the chip in extremely small, lightweight, and powerful laptop and notebook computers.

The 386 can execute the real-mode instructions of an 8086 or 8088, but in fewer clock cycles. The 386 was as efficient as the 286 in executing instructions, which means that the average instruction takes about 4.5 clock cycles. In raw performance, therefore, the 286 and 386 actually seemed to be about equal at equal clock rates. Many 286-system manufacturers were touting their 16 MHz and 20 MHz 286 systems as being just as fast as 16 MHz and 20 MHz 386 systems, and they were right! The 386 offered greater performance in other ways, mainly due to additional software capability (modes) and an incredibly enhanced Memory Management Unit (MMU).

The 386 can switch to and from protected mode under software control without a system reset, a capability that makes using protected mode more practical. In addition, the 386 has a new mode, called *virtual real mode*, which enables several real-mode sessions to run simultaneously under protected mode.

Other than raw speed, probably the most important feature of this chip is its available modes of operation, which are:

- Real mode
- Protected mode
- Virtual real mode (sometimes called *virtual 86 mode*)

Real mode on a 386 chip, as on a 286 chip, is 8086-compatible mode. In real mode, the 386 essentially is a much faster “turbo PC” with 640K of conventional memory, just like systems based on the 8088 chip. DOS and any software written to run under DOS requires this mode to run.

The protected mode of the 386 is fully compatible with the protected mode of the 286. The protected mode for both of these chips often is called their *native mode* of operation, because these chips are designed for advanced operating systems such as OS/2 and Windows NT, which run only in protected mode. Intel extended the memory-addressing capabilities of 386 protected mode with a new memory-management unit (MMU) that provides advanced memory paging and program switching. These features are extensions of the 286 type of MMU, so the 386 remains fully compatible with the 286 at system-code level.

The 386 chip’s virtual real mode is new. In virtual real mode, the processor can run with hardware memory protection while simulating an 8086’s real-mode operation. Multiple copies of DOS and other operating systems, therefore, can run simultaneously on this processor, each in a protected area of memory. If the programs in one segment crash, the rest of the system is protected. Software commands can reboot the blown partition.

In simple terms, a PC with a 386 has the capability to “become” multiple PCs under software control. With appropriate management software, the 386 chip can create several memory partitions, each containing the full services of DOS, and each of these partitions can function as though it were a stand-alone PC. These partitions often are called *virtual machines*.

In 386 virtual real mode under software such as Windows, several DOS programs can be running at the same time as programs designed for Windows. Because the processor can service only a single application at a time by delivering a clock tick, Windows manages the amount of CPU time that each program gets by using a system called *time slices*. Because the 386 chip is so fast and because time slices are tiny fractions of a second, under Windows, all applications appear to be running simultaneously.

OS/2 exploits the multitasking capabilities of the 386 chip even more than Windows does. OS/2 2.x can simultaneously manage native OS/2 programs, DOS programs, and most
Windows programs. These capabilities aren't available in lesser processors, such as the 286.

The 386 exploits protected mode much more effectively than the 286 does. The 386 can switch to and from protected mode under software control without a system reset. The 286 cannot switch from protected mode without a hardware reset.

Numerous variations of the 386 chip exist, some of which are less powerful and some of which are less power-hungry. The following sections cover the members of the 386-chip family and their differences.

386DX Processors

The 386DX chip was the first of the 386-family members that Intel introduced. The 386 is a full 32-bit processor with 32-bit internal registers, a 32-bit internal data bus, and a 32-bit external data bus. The 386 contains 275,000 transistors in a VLSI (Very Large Scale Integration) circuit. The chip comes in a 132-pin package and draws approximately 400 milliamperes (ma), which is less power than even the 8086 requires. The 386 has a smaller power requirement because it is made of CMOS (Complementary Metal Oxide Semiconductor) materials. The CMOS design enables devices to consume extremely low levels of power.

The Intel 386 chip is available in clock speeds ranging from 16 MHz to 33 MHz; other manufacturers offer comparable versions that offer speeds up to 40 MHz.

The 386DX can address 4G of physical memory. Its built-in virtual memory manager enables software designed to take advantage of enormous amounts of memory to act as though a system has 64 terabytes of memory. (A *terabyte* is 1,099,511,627,776 bytes of memory.) Although most 386 systems are built to accept 64M or less in RAM chips on the motherboard, some high-end computer users do take advantage of the 386 chip's capacity for 4G of physical memory, as well as its 64T virtual-memory potential.

386SX Processors

The 386SX, code-named the P9 chip during its development, was designed for systems designers who were looking for 386 capabilities at 286-system prices. Like the 286, the 386SX is restricted to only 16 bits when communicating with other system components. Internally, however, the 386SX is identical to the DX chip; the 386SX can handle 32 bits of data at the same time (compared with the 8 bits that the original IBM PC handled).

The 386SX uses a 24-bit memory-addressing scheme like that of the 286, rather than the full 32-bit memory address bus of the standard 386. The 386SX, therefore, can address a maximum 16M of physical memory rather than the 4G of physical memory that the 386DX can address. The 386SX is available in clock speeds ranging from 16 MHz to 33 MHz.

The 386SX signaled the end of the 286 because of the 386SX chip's superior MMU and the addition of the virtual real mode. Under a software manager such as Windows or OS/2, the 386SX can run numerous DOS programs at the same time. The capability to run 386-specific software is another important advantage of the 386SX over any 286 or older design. For example, Windows 3.1 runs nearly as well on a 386SX as it does on a 386DX.

Note

One common fallacy about the 386SX is that you can plug one into a 286 system and give the system 386 capabilities. This is not true; the 386SX chip is not pin-compatible with the 286 and does not plug into the same socket. Several upgrade products, however, have been designed to adapt the chip to a 286 system. In terms of raw speed, converting a 286 system to a 386 CPU chip results in little performance gain, because 286 motherboards are built with a restricted 16-bit interface to memory and peripherals. A 16 MHz 386SX is not markedly faster than a 16 MHz 286, but it does offer improved memory-management capabilities on a motherboard designed for it, as well as the capability to run 386-specific software.

386SL Processors

Another variation on the 386 chip is the 386SL. This low-power CPU has the same capabilities as the 386SX, but it is designed for laptop systems in which low power consumption is needed. The SL chips offer special power-management features that are important to systems that run on batteries. The SL chip offers several sleep modes that conserve power.

The chip includes an extended architecture that includes a System Management Interrupt (SMI), which provides access to the power-management features. Also included in the SL chip is special support for LIM (Lotus Intel Microsoft) expanded memory functions and a cache controller. The cache controller is designed to control a 16K-to-64K external processor cache.

These extra functions account for the higher transistor count in the SL chips (855,000) compared with even the 386DX processor (275,000). The 386SL is available in 25 MHz clock speed.

Intel offers a companion to the 386SL chip for laptops called the *82360SL I/O subsystem*. The 82360SL provides many common peripheral functions, such as serial and parallel ports, a direct memory access (DMA) controller, an interrupt controller, and power-management logic for the 386SL processor. This chip subsystem works with the processor to provide an ideal solution for the small size and low power-consumption requirements of portable and laptop systems.

386 Processor Clones

Several manufacturers, including AMD and Cyrix, have developed their own versions of the Intel 386DX and SX processors. These 386-compatible chips are available in speeds up to 40 MHz; Intel produces 386 chips up to only 33 MHz. Intel does not offer a 386 chip faster than 33 MHz because that speed begins to tread on the performance domain of the slowest of its own 486 processors.

In general, these chips are fully function-compatible with the Intel processors, which means that they run all software designed for the Intel 386. Many manufacturers choose these "cloned" 386 chips for their systems because they are faster and less expensive than Intel 386 chips. (Intel developed its "Intel Inside" advertising campaign in the hope of enticing buyers with a promise of getting the real thing.)

The section on IBM processors discusses the Intel-compatible chips designed and sold by IBM. These chips are not quite comparable with the other Intel processor clones, because they actually use official masks and microcode licensed directly from Intel. This arrangement essentially gives IBM the full design of the chip to use in its present form or modify. Thus, the IBM processors are fully compatible with the Intel processors and often offer many enhancements that are not even available in the Intel versions.

486 Processors

In the race for more speed, the Intel 80486 (normally abbreviated as 486) was another major leap forward. The additional power available in the 486 fueled tremendous growth in the software industry. Tens of millions of copies of Windows, and millions of copies of OS/2, have been sold largely because the 486 finally made the graphical user interface (GUI) of Windows and OS/2 a realistic option for people who work on their computers every day.

Three main features make a given 486 processor roughly twice as fast as an equivalent MHz 386 chip. These features are:

- *Reduced instruction-execution time*. Instructions in the 486 take an average of only 2 clock cycles to complete, compared with an average of more than 4 cycles on the 386.

- *Internal (Level 1) cache*. The built-in cache has a hit ratio of 90 percent to 95 percent, which describes how often zero-wait-state read operations will occur. External caches can improve this ratio further.

- *Burst-mode memory cycles*. A standard 32-bit (4-byte) memory transfer takes 2 clock cycles. After a standard 32-bit transfer, more data up to the next 12 bytes (or 3 transfers) can be transferred with only 1 cycle used for each 32-bit (4-byte) transfer. Thus, up to 16 bytes of contiguous, sequential memory data can be transferred in as little as 5 cycles instead of 8 cycles or more. This effect can be even greater when the transfers are only 8 bits or 16 bits each.

- *Built-in (synchronous) enhanced math coprocessor* (some versions). The math

> coprocessor runs synchronously with the main processor and executes math instructions in fewer cycles than previous designs do. On average, the math coprocessor built in to the DX-series chips provides two to three times greater math performance than an external 387 chip.

The 486 chip is about twice as fast as the 386, which means that a 386DX-40 is about as fast as a 486SX-20. If given a choice between a 40 MHz 386 and a 20 MHz 486, I would go for the 486. The lower-MHz 486 chip not only will be just as fast (or faster), but also can easily be upgraded to a DX2 or DX4 processor—which would be two or three times faster yet. You can see why the arrival of the 486 rapidly killed off the 386 in the marketplace.

Before the 486, many people avoided GUIs because they didn't have time to sit around waiting for the hourglass, which indicates that the system is performing behind-the-scenes operations that the user cannot interrupt. (It was said, only partly in jest, that you could turn on your system in the morning, start Windows or OS/2, and then go make coffee for the office. By the time you got back to your computer, Windows or OS/2 might be finished loading.) The 486 changed that situation. Many people believe that the 486 CPU chip spawned the widespread acceptance of GUIs.

The 486 chip's capability to handle the GUI prompted sales of pricey hardware: faster and larger hard drives, faster video display boards and larger monitors, faster and better printers, optical storage devices, CD-ROM drives, sound boards, and video capture boards. A fortunate occurrence prompted by all this spending is the fact that hardware (and software) prices have been in a steep decline for several years.

With the release of its faster Pentium CPU chip, Intel began to cut the price of the 486 line to entice the industry to shift over to the 486 as the mainstream system. Now Intel is starting to cut the price of the Pentium as well. The 486 chip is available in numerous versions: with and without math coprocessors, in clock speeds ranging from 16 MHz to 100 MHz, with special power-management capabilities, and with 3.3-volt operation to save even more power.

Besides high performance, one of the best features of the 486 family of chips is upgradability. In most cases, you can enjoy a performance increase in a given 486 system simply by adding or changing to a faster CPU. Unfortunately, Intel has not explained this upgradability well. I have found it difficult to cut through the marketspeak to find out the technical issues behind the different 486 CPU upgrades—specifically, how they work and what the ramifications of these upgrades are.

In this section, you will find information that can dispel any misunderstandings you may have about the 486 chip versions and possible upgrades. The section explains all the available 486 processors, as well as the possible upgrades and interchanges. You will learn the differences between items such as the new DX2 and Overdrive CPUs, and you will learn which items are appropriate for a given 486-class system. The sections that follow cover the variations on the basic 486 chip.

The 486 Processor Family. Since the introduction of the original 486DX chip in April 1989, the 486 has spawned an entire family of processors. Although 486 processors share certain features, such as full 32-bit architecture and a built-in memory cache, the various members of the 486 family differ in certain features, as well as in maximum speeds and pin configurations. This section first breaks down the different 486 processors by their major types, speed differences, and physical configuration, and then describes each processor in depth. Following are the current primary versions of the 486:

- 486SX—486 CPU without FPU (Floating-Point Unit)
- 486DX—486 CPU plus FPU
- 486DX2—Double-speed (Overdrive) 486 CPU plus FPU
- 486DX4—Triple-speed 486 CPU plus FPU

In addition, most of these 486 chips are available in a variety of maximum speed ratings, varying from 16 MHz at the low end to 100 MHz for the fastest chips. The following table shows the maximum speed ratings of the 486 processors.

Table 6.4 Standard Intel 486 Processor Maximum Clock Ratings

	Clock Speed in Megahertz (MHz)									
Processor Type	**16**	**20**	**25**	**33**	**40**	**50**	**66**	**75**	**83**	**100**
486SX	×	×	×	×						
486DX			×	×		×				
486DX2					×	×	×			
486DX4								×	×	×

A processor rated for a given speed always functions at any of the lower speeds. A 33 MHz-rated 486DX chip, for example, runs at 25 MHz if it is plugged into a 25 MHz motherboard. Notice that the DX2/Overdrive processors operate internally at 2 times the motherboard clock rate, whereas the DX4 processors operate at 2, 2-1/2, or 3 times the motherboard clock rate. The following table shows the different speed combinations that can result from using the DX2 or DX4 processors with different motherboard clock speeds.

Table 6.5 Intel DX2 and DX4 Operating Speeds versus Motherboard Clock Speeds

Motherboard Clock Speed	16 MHz	20 MHz	25 MHz	33 MHz	50 MHz
DX2 processor speed	32 MHz	40 MHz	50 MHz	66 MHz	N/A
DX4 (2x mode) speed	32 MHz	40 MHz	50 MHz	66 MHz	100 MHz

DX4 (2.5x mode) speed	40 MHz	50 MHz	62.5 MHz	83 MHz	N/A
DX4 (3x mode) speed	48 MHz	60 MHz	75 MHz	100 MHz	N/A

Besides differing in clock speeds, 486 processors have slight differences in pin configurations. The DX, DX2, and SX processors have a virtually identical 168-pin configuration, whereas the Overdrive chips sold retail have either the standard 168-pin configuration or a specially modified 169-pin Overdrive (sometimes also called 487SX) configuration. If your motherboard has two sockets, the primary one likely supports the standard 168-pin configuration, and the secondary (Overdrive) socket supports the 169-pin Overdrive configuration. Most newer motherboards with a single ZIF (zero insertion force) socket support any of the 486 processors except the DX4. The DX4 is different because it requires 3.3 volts to operate instead of 5 volts, like most of the other chips.

If you are upgrading an existing system, be sure that your socket will support the chip that you are installing. In particular, if you are putting one of the new DX4 processors in an older system, you need some type of adapter to regulate the voltage down to 3.3 volts. If you put the DX4 in a 5v socket, you will destroy the (very expensive) chip!

The 486-processor family is designed for high performance because it integrates formerly external devices such as cache controllers, cache memory, and math coprocessors. Also, 486 systems are designed for upgradability. Most 486 systems can be upgraded by simple processor additions or swaps that can effectively double the speed of the system. Because of these features, I recommend the 486SX or DX as the ideal entry-level system, especially in a business environment. Your investment will be protected in the future by a universally available, low-cost processor upgrade.

Internal (Level 1) Cache. All members of the 486 family include, as a standard feature, an integrated (Level 1) cache controller with either 8K or 16K of cache memory included. This cache basically is an area of very fast memory built into the processor that is used to hold some of the current working set of code and data. Cache memory can be accessed with no wait states, because it can fully keep up with the processor. Using cache memory reduces a traditional system bottleneck, because system RAM often is much slower than the CPU. This prevents the processor from having to wait for code and data from much slower main memory, therefore improving performance. Without the cache, a 486 frequently would be forced to wait until system memory caught up. If the data that the 486 chip wants is already in the internal cache, the CPU does not have to wait. If the data is not in the cache, the CPU must fetch it from the secondary processor cache or (in less sophisticated system designs) from the system bus.

You do not need special software or programs to take advantage of this cache; it works invisibly inside the chip. Because the cache stores both program instructions (code) and data, it is called a *unified cache*.

The organization of the cache memory in the 486 family technically is called a 4-Way Set

Associative Cache, which means that the cache memory is split into four blocks. Each block also is organized as 128 or 256 lines of 16 bytes each.

To understand how a 4-way set associative cache works, consider a simple example. In the simplest cache design, the cache is set up as a single block into which you can load the contents of a corresponding block of main memory. This procedure is similar to using a bookmark to locate the current page of a book that you are reading. If main memory equates to all the pages in the book, the bookmark indicates which pages are held in cache memory. This procedure works if the required data is located within the pages marked with the bookmark, but it does not work if you need to refer to a previously read page. In that case, the bookmark is of no use.

An alternative approach is to maintain multiple bookmarks to mark several parts of the book simultaneously. Additional hardware overhead is associated with having multiple bookmarks, and you also have to take time to check all the bookmarks to see which one marks the pages of data that you need. Each additional bookmark adds to the overhead but also increases your chance of finding the desired pages.

If you settle on marking four areas in the book to limit the overhead involved, you have essentially constructed a 4-way set associative cache. This technique splits the available cache memory into four blocks, each of which stores different lines of main memory. Multitasking environments, such as OS/2 and Windows, are good examples of environments in which the processor needs to operate on different areas of memory simultaneously and in which a four-way cache would improve performance greatly.

The contents of the cache must always be in sync with the contents of main memory to ensure that the processor is working with current data. For this reason, the internal cache in the 486 family is a Write-Through cache. *Write-Through* means that when the processor writes information out to the cache, that information is automatically written through to main memory as well.

By comparison, the Pentium chip has an internal *Write-Back* cache, which means that both reads and writes are cached, further improving performance. Even though the internal 486 cache is Write-Through, the system still can employ an external Write-Back cache for increased performance. In addition, the 486 can buffer up to 4 bytes before actually storing the data in RAM, improving efficiency in case the memory bus is busy.

The cache controller built into the processor also is responsible for watching the memory bus when alternate processors known as Bus Masters are in control of the system. This process of watching the bus is referred to as Bus Snooping. If a Bus Master device writes to an area of memory that also is stored in the processor cache currently, the cache contents and memory no longer agree. The cache controller then marks this data as invalid and reloads the cache during the next memory access, preserving the integrity of the system.

An external *secondary cache* (Level 2) of up to 512K or more of extremely fast Static RAM (SRAM) chips also is used in most 486-based systems to further reduce the amount of time that the CPU must spend waiting for data from system memory. The function of the secondary processor cache is similar to that of the 486 chip's on-board cache. The secondary processor cache holds information that is moving to the CPU, thereby reducing the time that the CPU spends waiting and increasing the time that the CPU spends performing calculations. Fetching information from the secondary processor cache rather than from system memory is much faster because of the extremely fast speed of the SRAM chips—20 nanoseconds (ns) or less.

The following sections discuss the technical specifications and differences of the various members of the 486-processor family in more detail.

486DX Processors. The original Intel 486DX processor was introduced on April 10, 1989, and systems using this chip first appeared during 1990. The first chips had a maximum speed rating of 25 MHz; later versions of the 486DX were available in 33 MHz- and 50 MHz-rated versions. The 486DX originally was available only in a 5v, 168-pin PGA (Pin Grid Array) version but now also is available in 5v, 196-pin PQFP (Plastic Quad Flat Pack) and 3.3v, 208-pin SQFP (Small Quad Flat Pack) as well. These latter form factors are available in SL Enhanced versions, which are intended primarily for portable or laptop applications in which saving power is important.

Two main features separate the 486 processor from older processors such as the 386 or 286: integration and upgradability. The 486DX integrates functions such as the math coprocessor, cache controller, and cache memory into the chip. The 486 also was designed with upgradability in mind; double-speed Overdrive are upgrades available for most systems.

The 486DX processor is fabricated with low-power CMOS (Complimentary Metal Oxide Semiconductor) technology. The chip has a 32-bit internal register size, a 32-bit external data bus, and a 32-bit address bus. These dimensions are equal to those of the 386DX processor. The internal register size is where the "32-bit" designation used in advertisements comes from. The 486DX chip contains 1.2 million transistors on a piece of silicon no larger than your thumbnail. This figure is more than four times the number of components on 386 processors and should give you a good indication of the 486 chip's relative power. The following table shows the technical specifications of the 486DX processor.

Table 6.6 Intel 486DX Processor

Specifications	
Introduced	April 10, 1989 (25 MHz) June 24, 1991 (50 MHz)
Maximum rated speeds	25, 33, 50 MHz
CPU clock multiplier	1x
Register size	32-bit
External data bus	32-bit
Memory address bus	32-bit

Maximum memory	4G
Integral-cache size	8K
Integral-cache type	4-way set associative, Write-Through
Burst-mode transfers	Yes
Number of transistors	1.2 million, 1.4 million (SL Enhanced models)
Circuit size	1 micron (25, 33 MHz), 0.8 micron (50 MHz and all SL Enhanced models)
External package	168-pin PGA, 196-pin PQFP*, 208-pin SQFP*
Math coprocessor	Integral Floating-Point Unit (FPU)
Power management	SMM (System Management Mode) in SL Enhanced models
Operating voltage	5v standard, 3.3v optional in 208-pin SQFP models

PGA = Pin Grid Array
PQFP = Plastic Quad Flat Pack
SQFP = Small Quad Flat Pack
**The PQFP and SQFP models are SL Enhanced only.*

The standard 486DX contains a processing unit, a Floating-Point Unit (math coprocessor), a memory-management unit, and a cache controller with 8K of internal-cache RAM.
Due to the internal cache and a more efficient internal processing unit, the 486 family of processors can execute individual instructions in an average of only 2 processor cycles. Compare this figure with the 286 and 386 families, both of which execute an average 4.5 cycles per instruction, or with the original 8086 and 8088 processors, which execute an average 12 cycles per instruction. At a given clock rate (MHz), therefore, a 486 processor is roughly twice as efficient as a 386 processor; a 16 MHz 486SX is roughly equal to a 33 MHz 386DX system, and a 20 MHz 486SX is equal to a 40 MHz 386DX system. Any of the faster 486s are way beyond the 386 in performance.

The 486 is fully instruction-set-compatible with previous Intel processors, such as the 386, but offers several additional instructions (most of which have to do with controlling the internal cache).

Like the 386DX, the 486 can address 4G of physical memory and manage as much as 64 terabytes of virtual memory. The 486 fully supports the three operating modes introduced in the 386: real mode, protected mode, and virtual real mode. In real mode, the 486 (like the 386) runs unmodified 8086-type software. In protected mode, the 486 (like the 386) offers sophisticated memory paging and program switching. In virtual real mode, the 486 (like the 386) can run multiple copies of DOS or other operating systems while simulating an 8086's real-mode operation. Under an operating system such as

Windows or OS/2, therefore, both 16-bit and 32-bit programs can run simultaneously on this processor with hardware memory protection. If one program crashes, the rest of the system is protected, and you can reboot the blown portion through various means depending on the operating software.

Built-In Math Coprocessor. The 486DX series has a built-in math coprocessor that sometimes is called an MCP (math coprocessor) or FPU (Floating-Point Unit). This series is unlike previous Intel CPU chips, which required you to add a math coprocessor if you needed faster calculations for complex mathematics. The FPU in the 486DX series is 100 percent software-compatible with the external 387 math coprocessor used with the 386, but it delivers more than twice the performance because it runs in synchronization with the main processor and executes most instructions in half as many cycles as the 386.

486SL. Intel originally announced a stand-alone chip called the 486SL. Now the SL as a separate chip has been discontinued, and all the SL enhancements and features are available in virtually all the 486 processors (SX, DX, and DX2) in what are called SL Enhanced versions. The stand-alone 486SL, therefore, has been discontinued. *SL Enhancement* refers to a special design that incorporates special power-saving features.

The SL Enhanced chips originally were designed to be installed in laptop or notebook systems that run on batteries, but they are finding their way into desktop systems as well. The SL Enhanced chips feature special power-management techniques, such as sleep mode and clock throttling, to reduce power consumption when necessary. These chips are available in 3.3v versions as well.

The following table shows the technical specifications of the 486SL processor.

Table 6.7 Intel 486SL Processor

Specifications	
Introduced	November 9, 1992
Maximum rated speeds	25, 33, 50 MHz
CPU clock multiplier	1x
Register size	32-bit
External data bus	32-bit
Memory address bus	32-bit
Maximum memory	4G
Integral-cache size	8K
Integral-cache type	4-Way Set Associative, Write-Through
Burst-mode transfers	Yes
Number of transistors	1.4 million
Circuit size	0.8 micron
External package	168-pin PGA, 196-pin PQFP, 208-pin SQFP, 227-pin LGA*
Math coprocessor	Integral Floating-Point Unit (FPU), optional in some models

Power management	System Management Mode (SMM)
Operating voltage	5v standard, 3.3v optional in 208-pin SQFP models

PGA = Pin Grid Array
PQFP = Plastic Quad Flat Pack
SQFP = Small Quad Flat Pack
LGA = Land Grid Array
**The LGA version has been discontinued.*

Intel has designed a power-management architecture called System Management Mode (SMM). This new mode of operation is totally isolated and independent from other CPU hardware and software. SMM provides hardware resources such as timers, registers, and other I/O logic that can control and power down mobile-computer components without interfering with any of the other system resources. SMM executes in a dedicated memory space called System Management Memory, which is not visible and does not interfere with operating-system and application software. SMM has an interrupt called System Management Interrupt (SMI), which services power-management events, and which is independent from and higher-priority than any of the other interrupts.

SMM provides power management with flexibility and security that were not available previously. For example, when an application program tries to access a peripheral device that is powered down for battery savings, a System Management Interrupt (SMI) occurs, powering up the peripheral device and reexecuting the I/O instruction automatically.

Intel also has designed a feature called suspend/resume in the SL processor. The system manufacturer can use this feature to provide the portable-computer user with instant-on-and-off capability. An SL system typically can resume (instant on) in one second from the suspend state (instant off) to exactly where it left off. You do not need to reboot, load the operating system, load the application program, and then load the application data. Simply push the suspend/resume button, and the system is ready to go.

The SL CPU was designed to consume almost no power in the suspend state. This feature means that the system can stay in the suspend state possibly for weeks and yet start up instantly right where it left off. While it is in the suspend state, an SL system can keep working data in normal RAM memory safe for a long time, but saving to a disk still is prudent.

486SX. The 486SX, introduced in April 1991, was designed to be sold as a lower-cost version of the 486. The 486SX is virtually identical to the full DX processor, but the chip does not incorporate the FPU or math coprocessor portion.

As you read earlier in this chapter, the 386SX was a scaled-down (some people would say crippled) 16-bit version of the full-blown 32-bit 386DX. The 386SX even had a completely different pinout and was not interchangeable with the more powerful DX version. The 486SX, however, is a different story. The 486SX is in fact a full-blown 32-bit 486 processor that is basically pin-compatible with the DX. A couple of pin functions are different or rearranged, but each pin fits into the same socket.

The 486SX chip is more a marketing quirk than new technology. Early versions of the 486SX chip actually were DX chips that showed defects in the math-coprocessor section. Instead of being scrapped, the chips simply were packaged with the FPU section disabled and sold as SX chips. This arrangement lasted for only a short time; thereafter, SX chips got their own mask, which is different from the DX mask. The transistor count dropped to 1.185 million (from 1.2 million) to reflect this new mask.

The 486SX chip is twice as fast as a 386DX with the same clock speed. Intel has marketed the 486SX as being the ideal chip for new computer buyers, because not much entry-level software uses the math-coprocessor functions. If you use software that *does* use or require the math coprocessor, you are well advised to stick with the DX series as a minimum.

The 486SX is available in 16, 20, 25, and 33 MHz-rated speeds, and it normally comes in a 168-pin version, although other surface-mount versions are available in SL Enhanced models.

The following table shows the technical specifications of the 486SX processor.

Table 6.8 Intel 486SX Processor

Specifications	
Introduced	April 22, 1991
Maximum rated speeds	16, 20, 25, 33 MHz
CPU clock multiplier	1x (2x in some SL Enhanced models)
Register size	32-bit
External data bus	32-bit
Memory address bus	32-bit
Maximum memory	4G
Integral-cache size	8K
Integral-cache type	4-Way Set Associative, Write-Through
Burst-mode transfers	Yes
Number of transistors	1.185 million, 1.4 million (SL Enhanced models)
Circuit size	1 micron, 0.8 micron (SL Enhanced models)
External package	168-pin PGA, 196-pin PQFP*, 208-pin SQFP*
Math coprocessor	None
Power management	SMM (System Management Mode) in SL Enhanced models
Operating voltage	5v standard, 3.3v optional in 208-pin SQFP models

PGA = Pin Grid Array
PQFP = Plastic Quad Flat Pack
SQFP = Small Quad Flat Pack

**The PQFP and SQFP models are SL Enhanced only.*

Despite what Intel's marketing and sales information implies, technically, no provision exists for adding a separate math coprocessor to a 486SX system; neither is a separate math coprocessor chip available to plug in. Instead, Intel wants you to add a new 486 processor with a built-in math unit and disable the SX CPU that already is on the motherboard. If this situation sounds confusing, read on, because this topic brings you to the most important aspect of 486 design: upgradability.

487SX. The 487SX math coprocessor, as Intel calls it, really is a complete 25 MHz 486DX CPU with an extra pin added and some other pins rearranged. When the 487SX is installed in the extra socket provided in a 486SX-CPU-based system, the 487SX turns off the existing 486SX via a new signal on one of the pins. The extra key pin actually carries no signal itself and exists only to prevent improper orientation when the chip is installed in a socket.

The 487SX takes over all CPU functions from the 486SX and also provides math coprocessor functionality in the system. At first glance, this setup seems rather strange and wasteful, so perhaps further explanation is in order. Fortunately, the 487SX turned out simply to be a stopgap measure while Intel prepared its real surprise: the Overdrive processor. The DX2/Overdrive speed-doubling chips, which are designed for the 487SX 169-pin socket, have the same pinout as the 487SX. These upgrade chips are installed in exactly the same way as the 487SX; therefore, any system that supports the 487SX also supports the DX2/Overdrive chips.

When the 486SX processor was introduced, Intel told motherboard designers to install an empty 169-pin socket in the motherboard for a 487SX math coprocessor and originally called this socket a Performance Upgrade Socket. At first, the only thing that you could plug into this socket was what Intel called a 487SX math coprocessor. The strange thing was that the 487SX was not really a math coprocessor at all, but a full 486DX processor!

The only difference between a 487SX and a 486DX is the fact that the 487SX uses the 169-pin rearranged pinout. When you plug a 487SX into the upgrade socket, a special signal pin that was not defined before (interestingly, *not* the extra 169th one) shuts down the original 486SX CPU, and the 487SX takes over. Because the 487SX functionally is a full-blown DX processor, you also get the math coprocessor functions that were left out of the original 486SX CPU. That is one of the reasons why the 487SX is so expensive; you really are buying more than you think. The real crime is that the original CPU sits silently in the system and does nothing!

Even though the 487SX basically is the same as the 486DX, you normally cannot install a "regular" 486DX processor in the Overdrive socket, because the pin designations are not the same. (I used the word *normally* because some motherboards have a jumper selection to allow for the different CPU pin configurations.) Because the 486SX actually uses a 168-pin design similar to that of the 486DX (even though it normally is installed in a 169-pin socket), you may be able to install a regular DX chip in the SX socket and have it work, but this capability depends somewhat on the flexibility of your motherboard.

Although in most cases you could upgrade a system by removing the 486SX CPU and replacing it with a 487SX (or even a DX or DX2/Overdrive). Intel originally discouraged this procedure and recommended that PC manufacturers include a dedicated upgrade (Overdrive) socket in their systems, because several risks were involved in removing the original CPU from a standard socket. (The following section elaborates on those risks.) Nowadays, Intel recommends—or even insists on—the use of a single processor socket of a ZIF (zero insertion force) design, which makes upgrading an easy task physically.

DX2/Overdrive Processors. On March 3, 1992, Intel introduced the DX2 speed-doubling processors. On May 26, 1992, Intel announced that the DX2 processors also would be available in a retail version called Overdrive. Originally, the Overdrive versions of the DX2 were available only in 169-pin versions, which meant that they could be used only with 486SX systems that had sockets configured to support the rearranged pin configuration.

On September 14, 1992, Intel introduced 168-pin Overdrive versions for upgrading 486DX systems. These new processors, which represent the ultimate in PC performance, are available in new systems and also can be added to existing 486 (SX or DX) systems as an upgrade, even if those systems do not support the 169-pin configuration. When you use this processor as an upgrade, you simply install the new chip in your system, which subsequently runs twice as fast. (As the guy in *RoboCop* said, “I *like* it!”)

The DX2/Overdrive processors run internally at twice the clock rate of the host system. If the motherboard clock is 25 MHz, for example, the DX2/Overdrive chip runs internally at 50 MHz; likewise, if the motherboard is a 33 MHz design, the DX2/Overdrive runs at 66 MHz. The DX2/Overdrive speed doubling has no effect on the rest of the system; all components on the motherboard run the same as they do with a standard 486 processor. Therefore, you do not have to change other components (such as memory) to accommodate the double-speed chip. In other words, you can achieve a significant performance gain simply by changing the CPU chip; you do not have to use faster (more expensive) motherboard circuitry.

The DX2/Overdrive chips are available in several speeds. Currently, three different speed-rated versions are offered:

- 40 MHz DX2/Overdrive for 16 MHz or 20 MHz systems
- 50 MHz DX2/Overdrive for 25 MHz systems
- 66 MHz DX2/Overdrive for 33 MHz systems

Notice that these ratings indicate the maximum speed at which the chip is capable of running. You could use a 66 MHz-rated chip in place of the 50 MHz- or 40 MHz-rated parts with no problem, although the chip will run only at the slower speeds. The actual speed of the chip is double the motherboard clock frequency. When the 40 MHz DX2/Overdrive chip is installed in a 16 MHz 486SX system, for example, the chip will function only at 32 MHz—exactly double the motherboard speed. Intel originally stated that no 100 MHz DX2/Overdrive chip will be available for 50 MHz systems—which technically has not been true since the DX4 can be set to run in a clock doubled mode and used in a 50 MHz motherboard (more information on that situation later in the chapter).

The only part of the DX2 chip that doesn't run at double speed is the *bus interface unit*, a region of the chip that handles I/O between the CPU and the outside world. By translating between the differing internal and external clock speeds, the bus interface unit makes speed doubling transparent to the rest of the system. The DX2 appears to the rest of the system to be a regular 486DX chip, but one that seems to execute instructions twice as fast.

DX2/Overdrive chips are based on the 0.8-micron circuit technology that was first used in the 50 MHz 486DX. The DX2 contains 1.1 million transistors in a three-layer form. The internal 8K cache, integer, and Floating-Point Units all run at double speed. External communication with the PC runs at normal speed to maintain compatibility.

The following table shows the technical specifications of the 486DX2/Overdrive processors.

Table 6.9 486DX2/Overdrive Processor Specifications

Introduced	March 3, 1992
Maximum rated speeds	40, 50, 66 MHz
CPU clock multiplier	2x
Register size	32-bit
External data bu	32-bit
Memory address bus	32-bit
Maximum memory	4G
Integral-cache size	8K
Integral-cache type	4-Way Set Associative, Write-Through
Burst-mode transfers	Yes
Number of transistors	1.1 million, 1.4 million (SL Enhanced models)
Circuit size	0.8 micron
External package	168-pin PGA, 169-pin PGA, 196-pin PQFP*, 208-pin SQFP*
Math coprocessor	Integral Floating-Point Unit (FPU)
Power management	SMM (System Management Mode) in SL Enhanced models
Operating voltage	5v standard, 3.3v optional in 208-pin SQFP models

PGA = Pin Grid Array
PQFP = Plastic Quad Flat Pack
SQFP = Small Quad Flat Pack
**The PQFP and SQFP models are SL Enhanced only.*

Besides upgrading existing systems, one of the best parts of the DX2 concept is the fact that system designers can introduce very fast systems by using cheaper motherboard designs, rather than the very costly designs that would support a straight high-speed clock. This means that a 50 MHz 486DX2 system is much less expensive than a straight 50 MHz 486DX system. In a 486DX-50 system, the system board operates at a true 50 MHz. In a 486DX2-50 system, the 486DX2 CPU operates internally at 50 MHz, but the motherboard operates at only 25 MHz.

You may be thinking that a true 50 MHz DX-processor-based system still would be faster than a speed-doubled 25 MHz system, and this generally is true, but the differences in speed actually are very slight—a real testament to the integration of the 486 processor and especially to the cache design.

When the processor has to go to system memory for data or instructions, for example, it has to do so at the slower motherboard operating frequency, such as 25 MHz. Because the 8K internal cache of the 486DX2 has a hit rate of 90 percent to 95 percent, however, the CPU has to access system memory only 5 percent to 10 percent of the time for memory reads. Therefore, the performance of the DX2 system can come very close to that of a true 50 MHz DX system and cost much less. Even though the motherboard runs only at 33.33 MHz, a system with a DX2 66 MHz processors ends up being faster than a true 50 MHz DX system, especially if the DX2 system has a good Level-2 cache.

Because the DX2 66 MHz chips are much cheaper than the straight 50 MHz DX models, the systems are cheaper as well. Additionally, the newer local buses operate best at 33 MHz and do not tolerate 50 MHz speeds without using buffering. All these factors have contributed to the elimination of true 50 MHz DX systems from most manufacturers' inventories.

Many 486 motherboard designs also include a secondary cache that is external to the cache integrated into the 486 chip. This external cache allows for much faster access when the 486 chip calls for external-memory access. The size of this external cache can vary anywhere from 16K to 512K or more. When you add a DX2 processor, an external cache is even more important for achieving the greatest performance gain, because this cache greatly reduces the wait states that the processor will have to add when writing to system memory or when a read causes an internal-cache miss. For this reason, some systems perform better with the DX2/Overdrive processors than others, usually depending on the size and efficiency of the external-memory cache system on the motherboard. Systems that have no external cache still will enjoy a near-doubling of CPU performance, but operations that involve a great deal of memory access will be slower.

At this writing, Intel has stated that it has no plans for a DX2/Overdrive chip for 50 MHz systems. Producing a speed-doubled Overdrive processor for 486DX-50 based systems would mean that the Overdrive chip would have to function internally at 100 MHz. Indirectly, Intel has solved this problem with the introduction of the DX4 processor.

Although the DX4 technically is not sold as a retail part, you can indeed purchase it from several vendors, along with the 3.3v voltage adapter that you need to install the chip in a 5v socket. These adapters have jumpers that enable you to select the DX4 clock multiplier and set it to 2x, 2.5x, or 3x mode. In a 50 MHz DX system, you could install a DX4/voltage-regulator combination set in 2x mode for a motherboard speed of 50 MHz and a processor speed of 100 MHz! Although you may not be able to take advantage of the latest local bus peripherals, you will, in any case, have one of the fastest 486-class PCs available.

Differences between DX2 and Overdrive Processors. One of the most common questions about the DX2/Overdrive processors is "What's the difference between a DX2 chip and an Overdrive chip?" Although the advertisements are somewhat misleading, the DX2 and Overdrive processor chips actually are the same thing. The real difference is the way that they are sold and the amenities that are (or are not included). The simple answer is that if the chip comes installed in a system, it's a DX2; if it comes in an Intel retail upgrade kit, it's an Overdrive processor.

Overdrive processors are DX2 chips that are sold as end-user-installable upgrades. Like math coprocessors, these processors are available at retail outlets and carry a limited lifetime warranty from Intel. Included with Overdrive processors are a user guide, a utilities disk, a chip-extractor tool and a grounding strap. Overdrive processors for 25 MHz and 33 MHz systems also include a heat sink, which already is attached to the top surface of the chip. Although not all systems have poor-enough air circulation to require the heat sink, its presence increases the number of systems that can use the Overdrive chips. Technical support direct from Intel is provided with each Overdrive processor.

DX2 chips are the raw CPUs, which are sold in quantity only to OEMs (Original Equipment Manufacturers), which install the chips in their systems as the primary microprocessors. The DX2 chips are sold in bulk and do not include the packaging, documentation, software utilities, extractor tool, and other items that are part of the retail package. DX2 chips also do not include a heat sink from Intel; it is up to the system manufacturer to determine whether a heat sink is needed for a particular application and to add one if it is needed.

The raw DX2 CPUs are classified by Intel as OEM products and are warranted only to the OEM or authorized Intel distributor for one year from the ship date. When that OEM or distributor sells the system or the CPU, that company extends the warranty to the purchaser. The warranty and support for a raw 486DX2 come from the company from which you purchase your system or the 486DX2 CPU, not from Intel.

"Vacancy." Perhaps you saw the Intel advertisements—both print and television—that featured a 486SX system with a neon Vacancy sign pointing to an empty socket next to the CPU chip. Unfortunately, these ads were not very informative, and they made it seem that only systems with the extra socket could be upgraded. When I first saw these ads, I was worried, because I had just purchased a 486DX system, and the advertisements implied that only 486SX systems with the empty Overdrive socket were upgradable. This, of course, was not true, but the Intel advertisements surely did not communicate that fact very well.

I later found out that upgradability does not depend on having an extra Overdrive socket in the system and that virtually any 486SX or DX system can be upgraded. The secondary Overdrive socket was designed simply to make upgrading easier and more convenient. What I mean is that even in systems that have the second socket, you can actually remove the primary SX or DX CPU and plug the Overdrive processor directly into the main CPU socket, rather than into the secondary Overdrive socket.

In that case, you would have an upgraded system with a single functioning CPU installed; you could remove the old CPU from the system and sell it or trade it in for a refund. Unfortunately, Intel does not offer a trade-in or core-charge policy; it simply does not want your old chip. For this reason, some people saw the Overdrive socket as being a way for Intel to sell more CPUs. Some valid reasons exist, however, to use the Overdrive socket and leave the original CPU installed.

One reason is that many PC manufacturers void the system warranty if the CPU has been removed from the system. Also, when systems are serviced, most manufacturers require that the system be returned with only the original parts; you must remove all add-in cards, memory modules, upgrade chips, and similar items before sending the system in for servicing. If you replace the original CPU when you install the upgrade, returning the system to its original condition will be much more difficult.

Another reason for using the upgrade socket is that if the main CPU socket is damaged when you remove the original CPU or install the upgrade processor, the system will not function. By contrast, if a secondary upgrade socket is damaged, the system still should work with the original CPU.

If you think that damaging the socket or chip are not valid concerns, you should know that it typically takes 100 pounds of insertion force to install a chip in a standard 169-pin screw machine socket. With this much force involved, you easily could damage either the chip or socket during the removal-and-reinstallation process.

Many motherboard manufacturers began using low-insertion-force (LIF) sockets, which typically require only 60 pounds of insertion force for a 169-pin chip. With the LIF or standard socket, I usually advise removing the motherboard so that you can support the board from behind when you insert the chip. Pressing down on the motherboard with 60 to 100 pounds of force can crack the board if it is not supported properly. A special tool also is required to remove a chip from one of these sockets.

Nowadays, nearly all motherboard manufacturers are using zero-insertion-force (ZIF) sockets. These sockets almost eliminate the risk involved in upgrading, because no insertion force is necessary to install the chip. Most ZIF sockets are handle-actuated; you simply lift the handle, drop the chip into the socket, and then close the handle. This design makes replacing the original processor with the upgrade processor an easy task. Because it is so simple to perform the upgrade with a ZIF socket, most motherboards that use such a socket have only one processor socket instead of two. This arrangement is a bonus: the unnecessary second socket does not waste the additional motherboard space, and you are forced to remove the otherwise-dormant original processor, which you then can sell or keep as a spare.

If a ZIF socket is not used, it is usually much easier to install an upgrade processor in an empty Overdrive socket than it is to remove the original CPU and then install the upgrade chip in the CPU socket. For these reasons, Intel now recommends that all 486 systems (SX and DX) use the two-socket approach or (more likely) use a single ZIF socket for the primary CPU as well as for any later upgrades.

Most single-socket systems can take any of the 486-family chips from the 486SX to the DX and the DX2/Overdrive. These boards usually have a set of jumpers or switches that enable you to select the type and speed of CPU that you are installing. In systems that have no second socket or ZIF socket, you may be more restricted in terms of the types of upgrades that you can install.

Some motherboards now include sockets beyond the original 169-pin Overdrive socket (now officially called Socket 1) for use in additional upgrades. These larger sockets can be both primary-CPU sockets and Overdrive sockets. This design not only accommodates the original 486DX or DX2 processor, but also allows for an upgrade to the next level of Overdrive CPU based on the Intel Pentium processors.

Overdrive Processors and Sockets

Intel has stated that all of their future processors will have Overdrive versions available for upgrading at a later date. As a result, Intel has developed a series of socket designs that will accommodate not only the original processor with which a system is shipped, but also the future Overdrive processor.

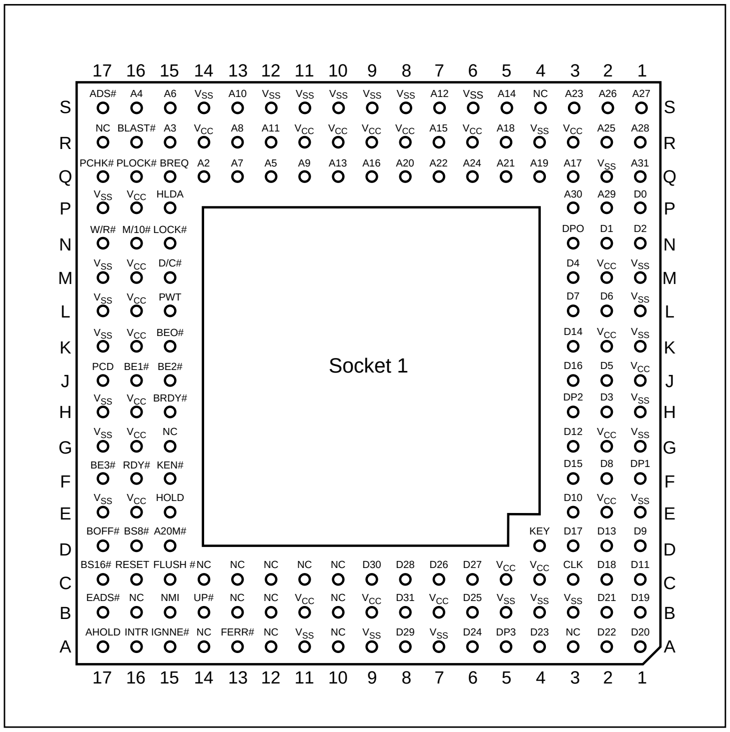
In many cases, the future Overdrive unit will be much more than just the same type of processor running at a higher clock rate. Although the original Overdrive series of processors for the 486SX and 486DX chip simply were clock-doubled versions of essentially the same chips, Intel plans Overdrive upgrades that go beyond this level. For example, the company already has designed Overdrive-style single-chip upgrades for DX2, DX4, and Pentium systems.

This new processor will require a larger socket than the original processor does; the additional pins are reserved for the new processor when it is ready. Intel is making available the pin specifications and some functions of the new processors so that motherboard designers can prepare now by installing the proper sockets. Then, when the Overdrive processor becomes available, all the end user will have to do is purchase it and install the new chip in place of the original one. To make the process easy, Intel now requires that all these sockets be of ZIF (zero insertion force) design.

Intel has created a set of six socket designs, named Socket 1 through Socket 6. Each socket is designed to support a different range of original and upgrade processors. The following table shows the specifications of these sockets.

Table 6.10 Intel 486/Pentium CPU Socket Types and Specifications

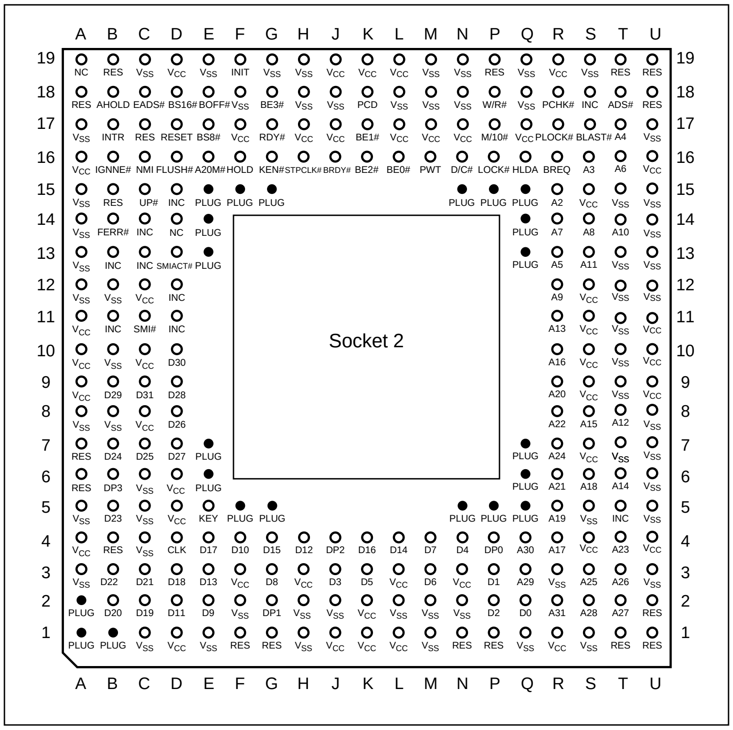
Socket	No. of	Pin



Number	Pins	Layout	Voltage	Supported Processors
Socket 1	169	17x17 PGA	5v	SX/SX2, DX/DX2*
Socket 2	238	19x19 PGA	5v	SX/SX2, DX/DX2*, Pentium Overdrive
Socket 3	237	19x19 PGA	5v/3.3v	SX/SX2, DX/DX2, DX4, Pentium Overdrive, DX4 Pentium Overdrive
Socket 4	273	21x21 PGA	5v	Pentium 60/66, Pentium 60/66 Overdrive
Socket 5	320	37x37 SPGA	3.3v	Pentium 90/100, Pentium 90/100 Overdrive
Socket 6	235	19x19 PGA	3.3v	DX4, DX4 Pentium Overdrive

**DX4 also can be supported with the addition of an aftermarket 3.3v voltage-regulator adapter.*
PGA = Pin Grid Array
SPGA = Staggered Pin Grid Array

The original Overdrive socket, now officially called Socket 1, is a 169-pin PGA socket. Motherboards that have this socket can support any of the 486SX, DX, and DX2 processors, as well as the DX2/Overdrive versions. This type of socket is found on most 486 systems that originally were designed for Overdrive upgrades. Even if your system has only a 168-pin version (technically, not Socket 1), you still can get DX2 chips with the correct pinout to plug right in. You even can install a DX4 triple-speed processor in this socket by using a voltage-regulator adapter. The following figure shows the pinout of



Socket 1.

Fig. 6.1
Intel Socket 1 pinout.

The original DX processor draws a maximum 0.9 amps of 5v power in 33 MHz form (4.5 watts) and a maximum 1 amp in 50 MHz form (5 watts). The DX2 processor or Overdrive processor draws a maximum 1.2 amps at 66 MHz (6 watts). This minor increase in power requires only a passive heat sink consisting of aluminum fins that are glued to the processor with thermal transfer epoxy. Overdrive processors rated at 40 MHz or less do not have heat sinks.

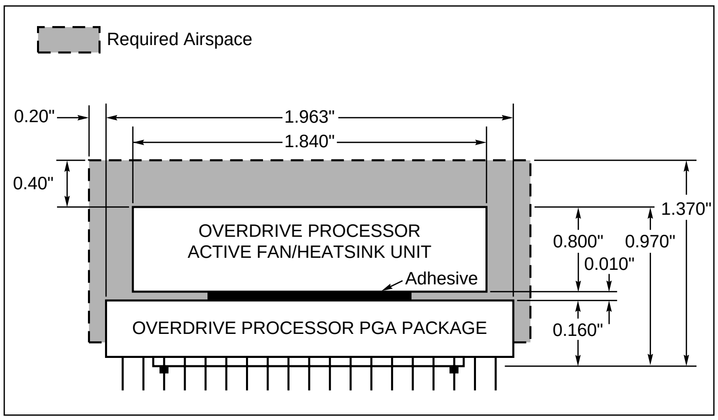
When the DX2 processor was released, Intel already was working on the new Pentium processor. The company wanted to offer a 32-bit, scaled-down version of the Pentium as an upgrade for systems that originally came with a DX2 processor. Rather than just increasing the clock rate, Intel created an all-new chip with enhanced capabilities derived from the Pentium.

The chip, officially called the Pentium Overdrive Processor, is not available at this writing but is scheduled to be available by the end of 1994, which may be when you read this. This chip will have 236 pins and will plug into a processor socket with the 238-pin Socket 2 design. This type of socket will hold any 486 SX, DX, or DX2 processor, as well as the Pentium Overdrive.

Figure 6.2 shows the pinout configuration of the official Socket 2 design.

Fig. 6.2
238-pin Intel Socket 2 configuration.

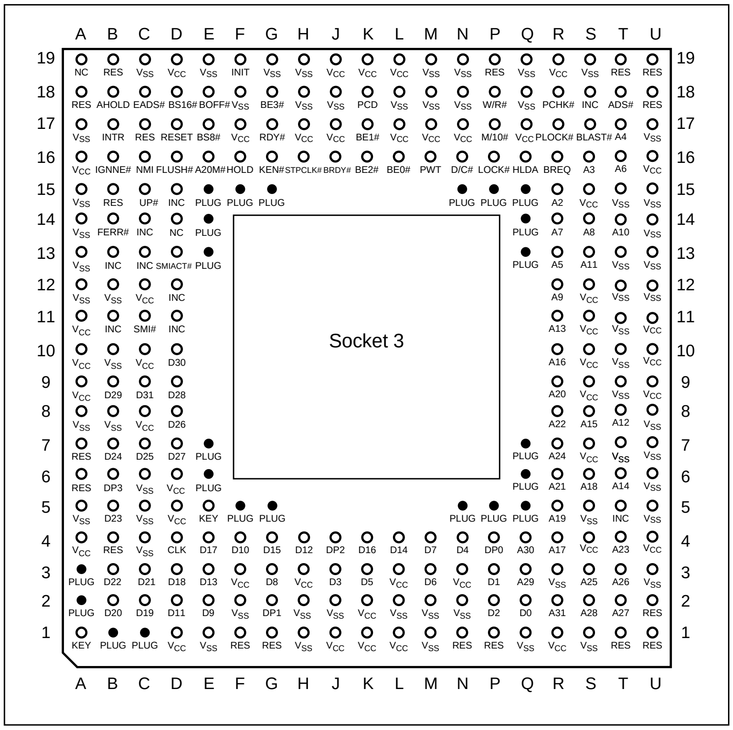
Notice that although the new chip for Socket 2 is called Pentium Overdrive, it is not a full-scale (64-bit) Pentium; the chip is more like a Pentium SX. Most manufacturers that claim to have Pentium-ready systems really mean that their systems have a Socket 2 processor socket that will accommodate the Pentium Overdrive chip in the future.



Intel released the design of Socket 2 a little prematurely and found that the chip ran too hot for many systems. The company solved this problem by adding a special Active Heat Sink to the Pentium Overdrive processor. This active heat sink is a combination of a standard heat sink with a built-in electric fan. Unlike the aftermarket glue-on or clip-on fans for processors that you may have seen, this one actually draws 5v power directly from the socket to drive the fan. No external connection to disk drive cables or the power supply is required. The fan/heat sink assembly clips or plugs directly into the processor, providing for easy replacement should the fan ever fail.

Another requirement of the active heat sink is additional clearance—no obstructions for an area about 1.4 inches off the base of the existing socket, to allow for heat-sink clearance. In systems that were not designed with this feature, the Pentium Overdrive upgrade will be difficult or impossible.

Another problem with this particular upgrade is power consumption. The 5v Pentium Overdrive processor will draw up to 2.5 amps at 5v (including the fan) or 12.5 watts, which is more than double the 1.2 amps (6 watts) drawn by the DX2 66 processor. Intel did not provide this information when it established the socket design, so the company



set up a testing facility to certify systems for thermal and mechanical compatibility with the Pentium Overdrive upgrade. For the greatest peace of mind, ensure that your system is certified compatible before you attempt this upgrade.

Figure 6.3 shows the dimensions of the Pentium Overdrive Processor and Active Heat Sink/fan assembly.

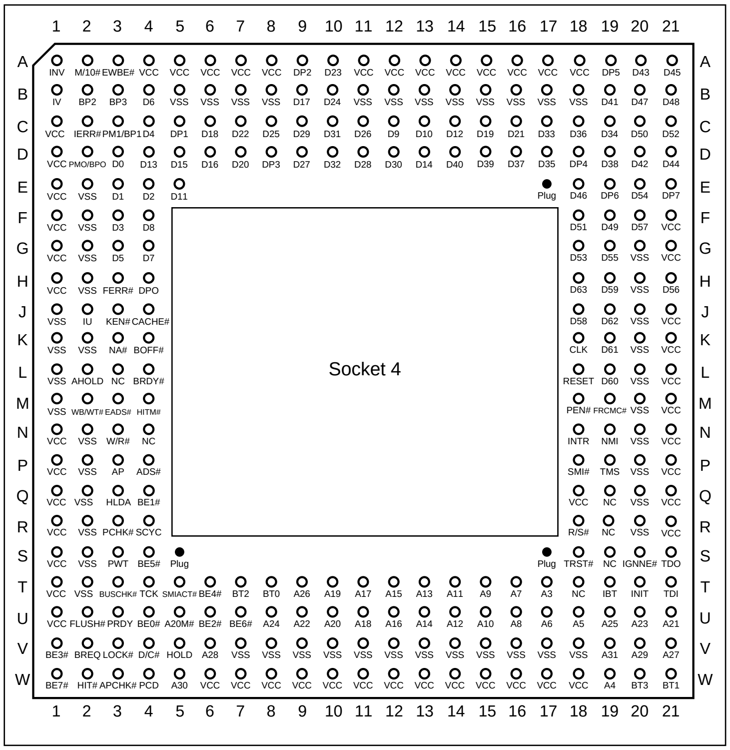
Fig. 6.3

Physical dimensions of the Intel Pentium Overdrive Processor and Active Heat Sink.

Because of problems with the original Socket 2 specification and the enormous heat the 5v version of the Pentium Overdrive Processor will generate, Intel came up with an improved design. The new processor, officially called the DX4 Overdrive Processor, actually is the same as the Pentium Overdrive processor, with the exception that it runs on 3.3v and draws a maximum 3.0 amps of 3.3v (9.9 watts) and 0.2 amp of 5v (1 watt) to run the fan, for a total 10.9 watts. This configuration provides a slight margin over the 5v version of this processor. The fan will be easy to remove from the Overdrive processor for replacement, should it ever fail.

To support both the new DX4 processor, which runs on 3.3v, and the 3.3v DX4 (Pentium) Overdrive processor, Intel had to create a new socket. In addition to the new

3.3v chips, this new socket supports the older 5v SX, DX, DX2, and even the 5v Pentium



Overdrive chip. The design, called Socket 3, is the most flexible upgradable 486 design.

Figure 6.4 shows the pinout specification of Socket 3.

Fig. 6.4
237-pin Intel Socket 3 configuration.

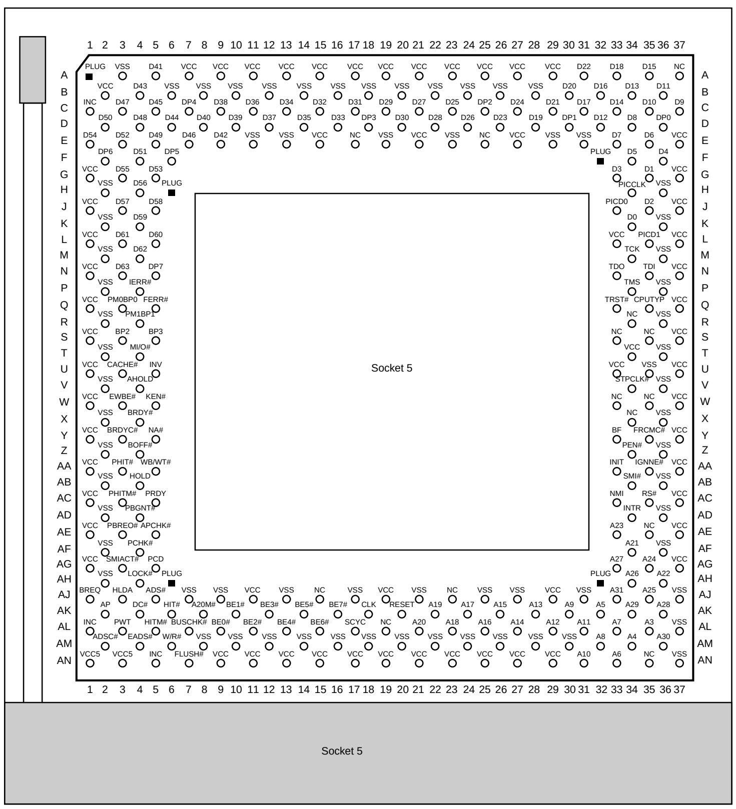
Notice that Socket 3 has one additional pin and several others plugged compared with Socket 2. Socket 3 provides for better keying, which prevents an end user from accidentally installing the processor in an improper orientation. One serious problem exists, however: this socket cannot automatically determine the type of voltage that will be provided to it. A jumper is likely to be added on the motherboard near the socket to enable the user to select 5v or 3.3v operation. Because this jumper must be manually set, however, a user could install a 3.3v processor in this socket when it is configured for 5v operation. This installation will instantly destroy a very expensive chip when the system is powered on. It will be up to the end user to make sure that this socket is properly configured for voltage, depending on which type of processor is installed. If the jumper is set in 3.3v configuration and a 5v processor is installed, no harm will occur, but the system will not operate properly unless the jumper is reset for 5v.

The original Pentium processor 60 MHz and 66 MHz versions had 273 pins and would plug into a 273-pin Pentium processor socket—a 5v-only socket, because all the original Pentium processors run on 5v. This socket will accept the original Pentium 60 MHz or 66 MHz processor, as well as the Overdrive processor.

Figure 6.5 shows the pinout specification of Socket 4.

Fig. *6.5*
273-pin Intel Socket 4 configuration.

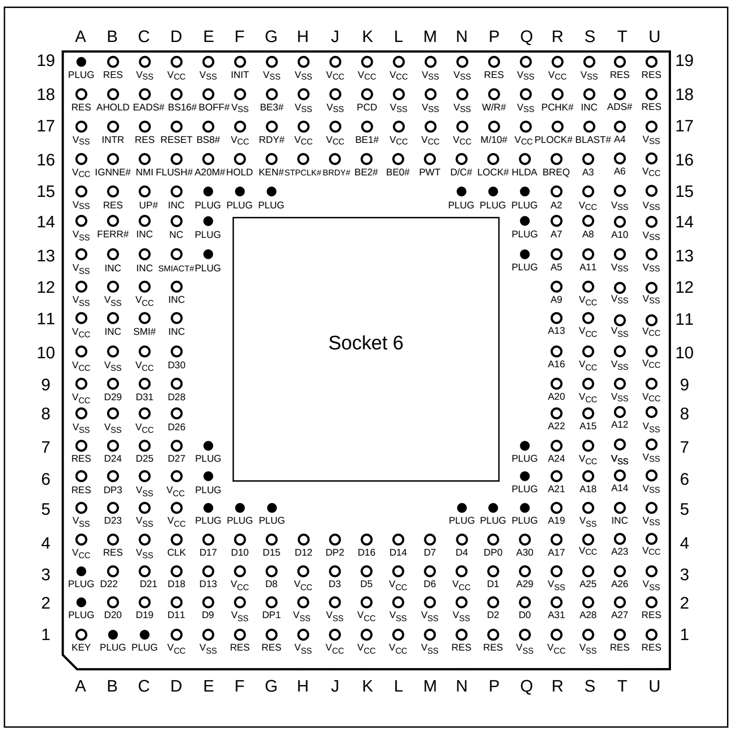
Somewhat amazingly, the original Pentium 66 MHz processor consumes up to 3.2 amps of 5v power (16 watts), not including power for a standard active heat sink (fan), whereas the 66 MHz Overdrive processor that will replace it consumes a maximum 2.7 amps (13.5 watts), including about 1 watt to drive the fan. Even the original 60 MHz Pentium processor consumes up to 2.91 amps at 5v (14.55 watts). It may seem strange that the replacement processor, which likely will be twice as fast, will consume less power than the original, but this has to do with the manufacturing processes used for the original and Overdrive processors.



Although both processors will run on 5v, the original Pentium processor was created with a circuit size of 0.8 micron, making that processor much more power-hungry than the newer 0.6-micron circuits used in the Overdrive and the other Pentium processors. Shrinking the circuit size is one of the best ways to decrease power consumption. Although the Overdrive processor for Pentium-based systems indeed will draw less power than the original processor, additional clearance may have to be allowed for the active heat sink (fan) assembly that is mounted on top. As in other Overdrive processors with built-in fans, the power to run the fan will be drawn directly from the chip socket, so no separate power-supply is required for connection. Also, the fan will be easy to replace, should it ever fail.

When Intel redesigned the Pentium processor to run at 90 and 100 MHz, the company went to a 0.6-micron manufacturing process as well as 3.3v operation. This change resulted in lower power consumption: only 3.25 amps at 3.3v (10.725 watts). Therefore, the 100 MHz Pentium processor can use far less power than even the original 60 MHz version.

The Pentium 90/100 processors actually have 296 pins, although they plug into the offi-



cial Intel Socket 5 design, which calls for a total 320 pins. The additional pins will be used by what officially is called the Future Pentium Overdrive Processor. This socket has the 320 pins configured in a Staggered Pin Grid Array, in which the individual pins are staggered for tighter clearance.

Figure 6.6 shows the standard pinout for Socket 5.

Fig. 6.6
320-pin Intel Socket 5 configuration.

The Future Pentium Overdrive Processor that eventually will use this socket will have an active heat sink (fan) assembly that will draw power directly from the chip socket. Intel has stated that this chip will require a maximum 4.33 amps of 3.3v to run the chip (14.289 watts) and 0.2 amp of 5v power to run the fan (1 watt), which means total power consumption of 15.289 watts. This amount is less power than the original 66 MHz Pentium processor requires, yet it runs a chip that is likely to be as much as four times faster!

The last socket is the newest design, which was created especially for the DX4 and the DX4 (Pentium) Overdrive Processor. Socket 6 basically is a slightly redesigned version of Socket 3, which has an additional two pins plugged for proper chip keying. Socket 6 has 235 pins and will accept only 3.3v 486 or Overdrive processors. Currently, this means that Socket 6 will accept only the DX4 and the DX4 (Pentium) Overdrive Processor. Because this socket provides only 3.3v, and because the only processors that plug into it are designed to operate on 3.3v, no chance exists that potentially damaging problems will occur, like those with the Socket 3 design. Most new 486-class systems that come with DX4 chips initially will use Socket 6. Figure 6.7 shows the Socket 6 pinout.

Fig. 6.7
235-pin Intel Socket 6 configuration.

The DX4 100 MHz processor can draw a maximum 1.45 amps of 3.3v (4.785 watts). The DX4 (Pentium) Overdrive Processor that eventually will replace that processor will draw a maximum 3.0 amps at 3.3v (9.9 watts) and 0.2 amp at 5v (1 watt) to run the fan, for a total 10.9 watts. Like all the other Overdrive processors that have active heat sink/fan assemblies, the processor has an easy-to-remove fan that can be replaced should it ever fail.

Overdrive Processor Installation. You can upgrade almost any system with an Overdrive processor. The most difficult aspect of the installation is simply having the correct Overdrive processor for your system. Currently, 486 Overdrive processors (which really are DX2 processors) are available for replacing 486SX and 486DX processors. The following table lists all the current and future Overdrive processors with their official Intel designations.

Processor Designation	Replaces	Socket	Heat Sink	Max. Power
486SX Overdrive	486SX	Socket 1	Passive	6 watts
486DX Overdrive	486DX	Socket 1	Passive	6 watts
486DX Overdrive	486DX	168-pin*	Passive	6 watts
Pentium Overdrive	486DX2	Socket 2 or 3	Active	12.5 watts

DX4 Overdrive	486DX4	Socket 3 or 6	Active	10.9 watts
Overdrive for Pentium	Pentium 60/66	Socket 4	Active	13.5 watts
Future Pentium Overdrive	Pentium 90/100	Socket 5	Active	15.289 watts

**The 168-pin version is designed for the original 486DX socket, which is slightly different from the official Socket 1 design.*

Three types of system configurations support the original 486 Overdrive (really DX2) processors:

- 486SX (Socket 1)
- 486DX (Socket 1)
- 486DX with standard 168-pin socket

Although Intel labels the Overdrive processors for SX or DX systems with a Socket 1 design, technically, the processors are exactly the same. The 168-pin version was specially created for systems that do not have the Socket 1 design, because it was created after the original 486DX systems were on the market. Many older 486DX systems will need this version of the Overdrive processor, which actually has the same pinout that is sold commercially as the DX2 processor, with the exception of the Intel-applied passive heat sink.

These Overdrive chips also are available in different speed ratings. You must select a version that is rated at least as fast as your motherboard will run. An Overdrive chip rated faster than is necessary will cost more but will work well; an Overdrive chip rated at a speed that is less than your motherboard will run will cause the chip to overheat and fail.

Most motherboards are speed-switchable, which means that if you currently have a 25 MHz 486SX system, you can upgrade to the 66 MHz Overdrive (or DX2) processor by first switching your motherboard to 33 MHz operation. If you leave the motherboard set at 25 MHz, the 66 MHz-rated chip would run at 50 MHz, but you would be missing out on potential performance.

To upgrade any 486SX or DX system that uses dual sockets—and therefore has a vacant Socket 1-type Overdrive Processor socket—you simply turn off the system, plug in the proper 169-pin Overdrive Processor, and turn the system back on. With the 169-pin versions designed for Socket 1, putting the processor in is literally impossible unless it is properly oriented with respect to pin 1. If you are using the 168-pin version that does not feature the key orientation pin, you must match the pin 1 indicator on the chip (usually a dot or a notched corner) with the appropriate designation on the socket (also a dot and/or notched corner). If you install the chip out of orientation, you likely will fry it when you power the system up.

To upgrade a system that has only a single processor socket, remove the existing processor and replace it with either a 169-pin (SX) or a 168-pin (DX) Overdrive Processor. If the system uses a ZIF (zero insertion force) socket, this procedure is very easy; otherwise, you must use a pry tool to remove the old chip.

Intel includes a pry-bar chip-extractor tool for removing the primary CPU, if that is necessary. You simply wedge the tool under one side of the chip and then pull to pry the chip partially out of the socket. You repeat this step for each of the chip's four sides. Then you can lift the loose chip out of the system and store it in a static-protective storage box (included with the Overdrive upgrade kit).

After you plug the Overdrive chip in, some systems require you to change jumper or switch settings on the motherboard to activate the chip. If you have an SX system, you also will have to run your system's setup program, because you must inform the CMOS memory that a math coprocessor is present. (Some DX systems also require you to run the setup program.) Intel provides a utility disk that includes a test program to verify that the new chip is installed and functioning correctly.

After verifying that the installation functions correctly, you have nothing more to do. You do not need to reconfigure any of the software on your system for the new chip. The only difference that you should notice is the fact that everything works nearly twice as fast as it did before the upgrade.

Upgrades that use the newer Overdrive chips for Sockets 2 through 6 are likely to be much easier, because these chips almost always go into a ZIF socket and therefore require no tools. In most cases, special configuration pins in the socket and on the new Overdrive chips take care of any jumper settings for you. In some cases, however, you may have
to set some jumpers on the motherboard to configure the socket for the new processor.

Overdrive Compatibility Problems. Although you can upgrade most older 486SX or 486DX systems with the Overdrive processors, some exceptions exist. Four factors can make an Overdrive upgrade difficult or impossible:

- BIOS routines that use CPU-dependent timing loops
- Lack of clearance for the Overdrive heat sink (25 MHz and faster)
- Inadequate system cooling
- A 486 CPU that is soldered in rather than socketed

To start with one of the obvious exceptions, you cannot use a DX2/Overdrive process to upgrade a 50 MHz 486DX system, because Intel does not make a DX2 Overdrive that is rated to run at the 100 MHz internal clock rate. In that case, you can perform an upgrade that is not officially authorized by Intel but that works in most cases: purchase a raw DX4 processor and voltage-regulator adapter to install in the 50 MHz 486DX socket. The voltage regulator will have a jumper for configuring the speed of the DX4 to 2x, 2.5x, or 3x operation. In this case, you need to set the jumper to 2x for 100 MHz operation, which is the maximum for which the DX4 100 chip is rated.

Because this upgrade is not officially authorized by Intel, and because Intel does not even sell the DX4 chips through retail channels, Intel provides no warranty. Make sure that the vendor from which you purchase the DX4 chip will take it back in case it does not work.

Systems that come with DX2, DX4, or Pentium chips installed by the system manufacturer also are eligible for future Overdrive upgrades, which may not yet be available.

In some rare cases, problems may occur in systems that should be upgradable but are not. One of these problems is related to the ROM BIOS. A few 486 systems have a BIOS that regulates hardware operations by using timing loops, based on how long it takes the CPU to execute a series of instructions. When the CPU suddenly is running twice as fast, the prescribed timing interval is too short, resulting in improper system operation or even hardware lockups. Fortunately, you usually can solve this problem by upgrading the system's BIOS.

Another problem is related to physical clearance. All Overdrive chips for 25 MHz and faster systems have heat sinks glued or fastened to the top of the chip. The heat sink can add 0.25 inch to 1.2 inches to the top of the chip. This extra height can interfere with other components in the system, especially in small desktop systems and portables. Solutions to this problem must be determined on a case-by-case basis. Sometimes, you can relocate an expansion card or disk drive, or even modify the chassis slightly to increase clearance. In some cases, the interference cannot be resolved, leaving you only the option of running the chip without the heat sink. Needless to say, removing the glued-on heat sink will at best void the warranty provided by Intel and will at worst damage the chip or the system due to overheating. I do not recommend removing the heat sink.

The Overdrive chips can generate up to twice the heat of the chips that they replace. Even with the active heat sink/fan built in to the faster Overdrive chips, some systems do not have enough airflow or cooling capability to keep the Overdrive chip within the prescribed safe operating-temperature range. Small desktop systems or portables are most likely to have cooling problems. Unfortunately, only proper testing can indicate whether a system will have a heat problem. For this reason, Intel has been running an extensive test program to certify systems that are properly designed to handle an Overdrive upgrade.

Finally, some systems have the 486SX or DX chip soldered directly into the motherboard rather than in a socket. This method is used sometimes for cost reasons, because leaving out the socket is cheaper; in most cases, however, the reason is clearance. The IBM P75 portable that I use, for example, has a credit-card-size CPU board that plugs into the motherboard. Because the CPU card is close to one of the expansion slots, to allow for clearance between the 486 chip and heat sink, IBM soldered the CPU directly into the
small card, making an Overdrive upgrade nearly impossible unless IBM offers its own

upgrade via a new CPU card with the DX2 chip already installed. This situation did not stop me, of course; I desoldered the DX chip and soldered in a pin-compatible 168-pin DX2 processor in its place. I currently am installing a voltage regulator and DX4 processor.

To clarify which systems are tested to be upgradable without problems, Intel has compiled an extensive list of compatible systems. To determine whether a PC is upgradable with an Overdrive processor, contact Intel via its FaxBACK system (see the vendor list in Appendix B) and ask for the Overdrive Processor Compatibility Data documents. These documents list the systems that have been tested with the Overdrive processors and indicate which other changes you may have to make for the upgrade to work (a newer ROM BIOS or setup program, for example).

One important note about these compatibility lists: if your system is not on the list, the warranty on the Overdrive processor is void. Intel recommends Overdrive upgrades only for systems that are in the compatibility list. The list also includes notes about systems that may require a ROM upgrade, a jumper change, or perhaps a new setup disk.

Some IBM PS/2 486 systems, for example, may require you to use a new Reference Diskette when you install an Overdrive CPU. You can download the latest version of any PS/2 Reference Diskette from the IBM National Support Center Bulletin Board System (NSC BBS), which appears in the vendor list. A wise practice is to download the latest version of the Reference and Diagnostics diskettes for any PS/2 system before attempting a processor upgrade.

Notice that the files on the IBM NSC BBS are compressed Reference and Diagnostics Diskette images. To decompress and extract the files in the appropriate format, you need one or both of the following files:

FILENAME.EXT	Contents
LDF.COM	".DSK" file-extraction program
TGSFX.COM	".TG0" file-extraction program

These files contain programs that will create the Reference or Diagnostics Diskette from the compressed ".DSK" or ".TG0" files.

Pentium Overdrive for DX2 and DX4 Systems. In late 1994, the Pentium Overdrive Processor (code-named P24T) and the DX4 (Pentium) Overdrive Processor will be released. These chips will be virtually identical, except that the first one will run on 5v, whereas the DX4 version will run on 3.3v and consume slightly less power.

The Pentium Overdrive Processor will be required for systems that have a processor socket that follows the Intel Socket 2 specification. This processor also will work in systems that have a Socket 3 design, although the DX4 3.3v version (the DX4 Overdrive Processor) is much more desirable. Just make sure that if you are using a 3.3v processor, you have Socket 3 set for 3.3v operation. Plugging the 3.3v version of the Overdrive chip into the Socket 2 design will be impossible; special key pins will prevent improper insertion. If your system has a 3.3v-only Socket 6 design, the only Overdrive processor to get is the DX4 Overdrive Processor.

Besides a 32-bit Pentium core, these processors will feature increased clock-speed operation, due to internal clock multiplication, and will incorporate an internal Write-Back cache (standard with the Pentium). In essence, these Overdrive processors are Pentium SX chips, because they will have all the features of the real Pentium processors except the 64-bit external data bus; the external data bus of these Overdrive chips will be 32 bits instead. Even so, with the improved Pentium core, separate code and Write-Back data caches, higher clock speeds and other enhancements, these processors should nearly double the performance of the systems in which they are installed.

If you are purchasing a 486 motherboard with the idea of performing this Pentium-level upgrade in the future, make sure that your motherboard has a Socket 3 or Socket 6 design so that you can take advantage of the improved 3.3v version of the Pentium Overdrive Processor.

Pentium

On October 19, 1992, Intel announced that the fifth generation of its compatible microprocessor line (code-named P5) would be named the Pentium processor rather than the 586, as everybody had been assuming. Calling the new chip the 586 would have been natural, but Intel discovered that it could not trademark a number designation, and the company wanted to prevent other manufacturers from using the same name for any clone chips that they might develop.

The actual Pentium chip shipped on March 22, 1993. Systems that use these chips were only a few months behind.

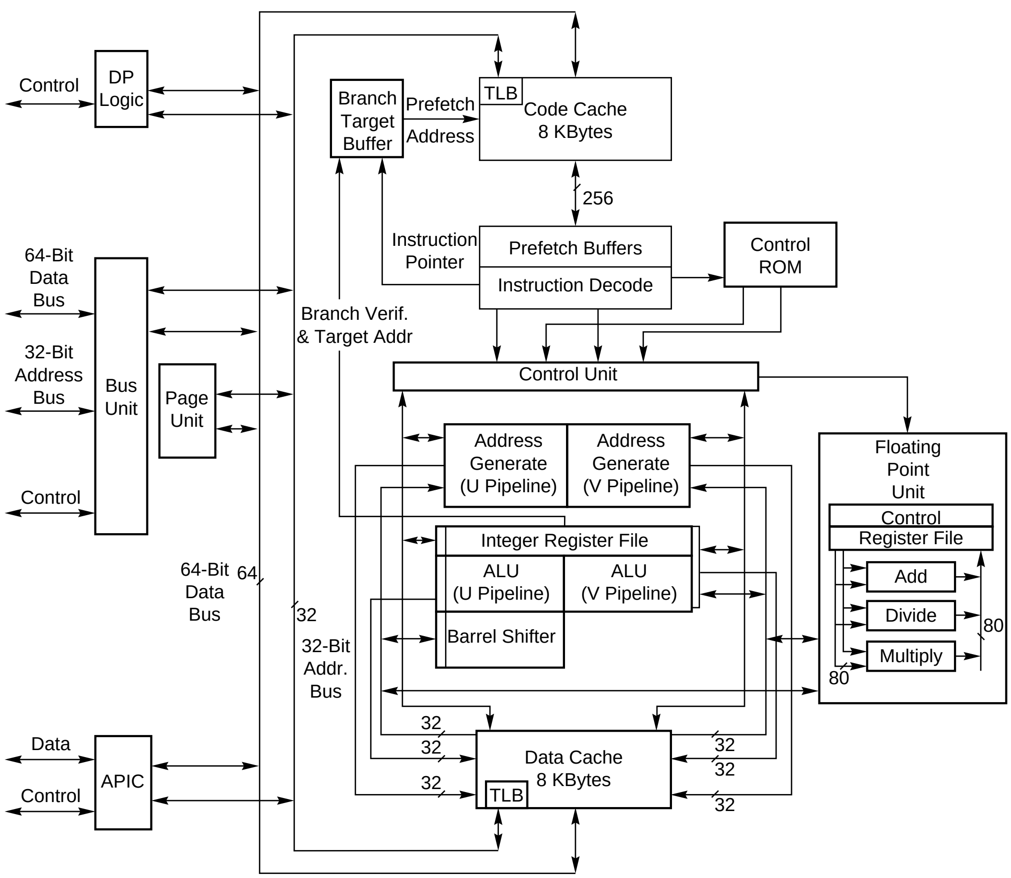
The Pentium is fully compatible with previous Intel processors, but it also differs from them in many ways. At least one of these differences is revolutionary: the Pentium features twin data pipelines, which enable it to execute two instructions at the same time. The 486 and all preceding chips can perform only a single instruction at a time. Intel calls the capability to execute two instructions at the same time *superscalar* technology. This technology provides additional performance compared with the 486.

The standard 486 chip can execute a single instruction in an average two clock cycles—cut to an average of one clock cycle with the advent of internal clock multiplication used in the DX2 and DX4 processors. With superscalar technology, the Pentium can execute many instructions at a rate of two instructions per cycle. Superscalar architecture usually is associated with high-output RISC (Reduced Instruction Set Computer) chips. The Pentium is one of the first CISC (Complex Instruction Set Computer) chips to be considered to be superscalar. The Pentium is almost like having two 486 chips under the hood. Table 6.11 shows the Pentium processor specifications.

Table 6.11 Pentium Processor

Specifications	
Introduced	March 22, 1993 (first generation); March 7, 1994 (second

	generation)
Maximum rated speeds	60, 66 MHz (first generation); 75, 90, 100 MHz (second generation)
CPU clock multiplier	1x (first generation), 1.5x–2x (second generation)
Register size	32-bit
External data bus	64-bit
Memory address bus	32-bit
Maximum memory	4G
Integral-cache size	8K code, 8K data
Integral-cache type	2-Way Set Associative, Write-Back Data
Burst-mode transfers	Yes
Number of transistors	3.1 million (60/66 MHz), 3.3 million (75 MHz and up)
Circuit size	0.8 micron (60/66 MHz), 0.6 micron (75 MHz and up)



External package	273-pin PGA, 296-pin SPGA
Math coprocessor	Built-in FPU (Floating-Point Unit)
Power management	SMM (System Management Mode), enhanced in second generation

Operating voltage	5v (first generation), 3.3v (second generation)

PGA = Pin Grid Array
SPGA = Staggered Pin Grid Array

The two instruction pipelines within the chip are called the *u- and v-pipes*. The u-pipe, which is the primary pipe, can execute all integer and floating-point instructions. The v-pipe is a secondary pipe that can execute only simple integer instructions and certain floating-point instructions. The process of operating on two instructions simultaneously in the different pipes is called *pairing*. Not all sequentially executing instructions can be paired, and when pairing is not possible, only the u-pipe is used. To optimize the Pentium's efficiency, you can recompile software to allow more instructions to be paired.

The Pentium is 100 percent software-compatible with the 386 and 486, and although all current software will run much faster on the Pentium, many software manufacturers want to recompile their applications to exploit even more of the Pentium's true power. Intel has developed new compilers that will take full advantage of the chip; the company will license the technology to compiler firms so that software developers can take advantage of the superscalar (parallel processing) capability of the Pentium. This optimization is starting to appear in some of the newest software on the market. Optimized software should improve performance by allowing more instructions to execute simultaneously in both pipes.

To minimize stalls in one or more of the pipes, caused by delays in fetching instructions that branch to nonlinear memory locations, the Pentium processor has a Branch Target Buffer (BTB) that employs a technique called *branch prediction*. The BTB attempts to predict whether a program branch will be taken or not and then fetches the appropriate next instructions. The use of branch prediction enables the Pentium to keep both pipelines operating at full speed. The following figure shows the internal architecture of the Pentium processor.

Fig. 6.8
Pentium processor internal architecture.

The Pentium has a 32-bit address bus width, giving it the same 4G memory-addressing capabilities as the 386DX and 486 processors. But the Pentium expands the data bus to 64 bits, which means that it can move twice as much data into or out of the CPU compared with a 486 of the same clock speed. The 64-bit data bus requires that system memory be accessed 64 bits wide, which means that each bank of memory is 64 bits.

On most motherboards, memory is installed via SIMMs (Single In-Line Memory Modules), and SIMMs are available in 9-bit-wide and 36-bit-wide versions. Most Pentium systems use the 36-bit-wide (32 data bits plus 4 parity bits) SIMMs—4 of these SIMMs per bank of memory. Most Pentium motherboards have 4 of these 36-bit SIMM sockets, providing for a total two banks of memory.

Even though the Pentium has a 64-bit data bus that transfers information 64 bits at a time into and out of the processor, the Pentium has only 32-bit internal registers. As instructions are being processed internally, they are broken down into 32-bit instructions and data elements, and processed in much the same way as in the 486. Some people thought that Intel was misleading them by calling the Pentium a 64-bit processor, but 64-bit transfers do indeed take place. Internally, however, the Pentium has 32-bit registers that are fully compatible with the 486.

The Pentium has two separate internal 8K caches, compared with a single 8K or 16K cache in the 486. The cache-controller circuitry and the cache memory are embedded in the CPU chip. The cache mirrors the information in normal RAM by keeping a copy of the data and code from different memory locations. The Pentium cache also can hold information to be written to memory when the load on the CPU and other system components is less. (The 486 makes all memory writes immediately.)

The separate code and data caches are organized in a two-way set associative fashion, with each set split into lines of 32 bytes each. Each cache has a dedicated Translation Lookaside Buffer (TLB), which translates linear addresses to physical addresses. You can configure the data cache as Write-Back or Write-Through on a line-by-line basis. When you use the Write-Back capability, the cache can store write operations as well as reads, further improving performance over read-only Write-Through mode. Using Write-Back mode results in less activity between the CPU and system memory—an important improvement, because CPU access to system memory is a bottleneck on fast systems. The code cache is an inherently write-protected cache, because it contains only execution instructions and not data, which is updated. Because burst cycles are used, the cache data can be read or written very quickly.

Systems based on the Pentium can benefit greatly from *secondary processor caches* (Level 2), which usually consist of up to 512K or more of extremely fast (20 ns or less) Static RAM (SRAM) chips. When the CPU fetches data that is not already available in its internal processor (Level 1) cache, wait states slow the CPU. If the data already is in the secondary processor cache, however, the CPU can go ahead with its work without pausing for wait states.

The Pentium uses a BiCMOS (Bipolar Complementary Metal Oxide Semiconductor) process and superscalar architecture to achieve the level of performance expected from the new chip. BiCMOS adds about 10 percent to the complexity of the chip design, but adds about 30 to 35 percent better performance without a size or power penalty. CMOS designs are running out of bandwidth at 66 MHz, and although CMOS can be made faster, a BiCMOS design provides room to go to 100 or 150 MHz and beyond. Intel probably will use the BiCMOS process in all future chip generations beyond the Pentium.

All Pentium processors are SL Enhanced, meaning that they incorporate the System Management Mode (SMM) to provide full control of power-management features, which helps reduce power consumption. The second-generation Pentium processors (75 MHz and faster) incorporate a more advanced form of SMM that includes processor clock control, which allows you to throttle the processor up or down to control power use. With these more advanced Pentium processors, you even can stop the clock, putting the processor in a state of suspension that requires very little power. The second-generation Pentium processors run on 3.3v power (instead of 5v), reducing power requirements and heat generation even further.

The Pentium, like the 486, contains an internal math coprocessor or Floating-Point Unit (FPU). The FPU in the Pentium has been rewritten and performs significantly better than the FPU in the 486, yet it is fully compatible with the 486 and 387 math coprocessor. The Pentium FPU is estimated to be 2 to as much as 10 times faster than the FPU in the 486. In addition, the two standard instruction pipelines in the Pentium provide two units to handle standard integer math. (The math coprocessor handles only more complex calculations.) Other processors, such as the 486, have only a single standard execution pipe and one integer-math unit.

First-Generation Pentium Processor. Currently, the Pentium is available in two designs, each with several versions. The first-generation design is available in 60 and 66 MHz processor speeds. This design uses a 273-pin PGA form factor and runs on 5v power. In this design, the processor runs at the same speed as the motherboard—in other words, a 1x clock is used.

The first-generation Pentium was created through an 0.8-micron BiCMOS process. Unfortunately, this process, combined with the 3.1 million transistor count, resulted in a die that was overly large and complicated to manufacture. As a result, reduced yields kept the chip in short supply; Intel could not make them fast enough. The 0.8-micron process was criticized by other manufacturers, including Motorola and IBM, which had been using 0.6-micron technology for their most advanced chips. The huge die and 5v operating voltage caused the 66 MHz versions to consume up to an incredible 3.2 amps or 16 watts of power, resulting in a tremendous amount of heat—and problems in some systems that did not employ conservative design techniques. Often, the system required a separate fan to blow on the processor, keeping it cool.

Much of the criticism leveled at Intel for the first-generation Pentium was justified. Some people realized that the first-generation design was just that; they knew that new Pentium versions, made in a more advanced manufacturing process, were coming. Many of those people (including the author of this book) advised against purchasing any Pentium system until the second-generation version became available.

A cardinal rule of computing is never to buy the first generation of any processor. Although you can wait forever, because something better always will be on the horizon, a little waiting is worthwhile in some cases.

Second-Generation Pentium Processor. Intel announced the second-generation Pentium on March 7, 1994. This new processor was introduced in 90 and 100 MHz versions, with a 75 MHz version for laptop and portable systems in development. The second-generation Pentium uses 0.6-micron BiCMOS technology to shrink the die and reduce power consumption. Additionally, these new processors run on 3.3v power. The 100 MHz version consumes a maximum 3.25 amps of 3.3v power, which equals only 10.725 watts. The slower 90 MHz version uses only 2.95 amps of 3.3v power, which is only 9.735 amps. The 75 MHz version probably will use about 6 watts of power and will function reasonably well in laptop and portable systems that run on batteries.

The second-generation Pentium processors come in a 296-pin SPGA (Staggered Pin Grid Array) form factor that is physically incompatible with the first-generation versions. The only way to upgrade from the first generation to the second is to replace the motherboard. The second-generation Pentium processors also have 3.3 million transistors—more than the earlier chips. The extra transistors exist because additional clock-control SL enhancements were added, as were an on-chip Advanced Programmable Interrupt Controller (APIC) and dual processor interface.

The APIC and dual processor interface are responsible for orchestrating dual-processor configurations in which two second-generation Pentium chips can process on the same motherboard simultaneously. Many of the new Pentium motherboards will come with dual Socket 5 specification sockets, which fully support the multiprocessing capability of the new chips. Already, software support for what usually is called Symmetric Multi-Processing (SMP) is being integrated into operating systems such as Windows and OS/2.

The second-generation Pentium processors use clock-multiplier circuitry to run the processor at speeds faster than the bus. The 90 MHz Pentium processor can run at 1.5 times the bus frequency, which normally is 60 MHz. The 100 MHz Pentium processor can run at a 1.5x clock in a system using a 66 MHz bus speed or at 2x clock on a motherboard that is running at 50 MHz. The future 75 MHz version reportedly also will use the 1.5x clock and, therefore, run on 50 MHz motherboards.

Currently, running the motherboard faster than 66 MHz is impractical because of memory and local-bus performance constraints. The fastest Pentium systems would combine the 66 MHz motherboard operation with a 1.5x clock and 100 MHz processor operation. If you think that 1.5 times 66 equals 99 and not 100, you may be right, but in nearly all cases, 66 MHz operation really means 66.6666 MHz actual speed.

Now that the second-generation Pentium processors have arrived, the time is right to purchase a Pentium system. The ideal system uses the second-generation 100 MHz processor with a 66 MHz motherboard.

Make sure that your Pentium motherboard includes two 320-pin processor sockets that fully meet the Intel 320-pin Socket 5 specification. These sockets enable you to add a second Pentium processor to take advantage of SMP (Symmetric Multi-Processing) support in the newer operating systems.

Also make sure that the motherboard can be jumpered or somehow reconfigured for 66 MHz operation, which will enable you to take advantage of future Pentium Overdrive processors that will support the higher motherboard clock speeds.

These simple recommendations will enable you to perform several dramatic upgrades without changing the entire motherboard.

Table 6.12 IBM Processor Specifications

Processor	CPU Clock	Std. Voltage	Internal Register Size	Data Bus Width	Address Bus Width
386 SLC	1x	5v	32-bit	16-bit	24-bit
486 SLC	1x	3.3v	32-bit	16-bit	24-bit
486 SLC2	2x	3.3v	32-bit	16-bit	24-bit
486 BL2	2x	3.3v	32-bit	32-bit	32-bit
486 BL3	3x	3.3v	32-bit	32-bit	32-bit

FPU = Floating-Point Unit (math coprocessor)

IBM (Intel-Licensed) Processors

For many years, IBM and Intel have had a very close relationship, especially considering the fact that IBM is one of Intel's largest customers. The companies often enter into agreements to exchange technology and information. One such agreement involves the licensing of Intel's processors; IBM has a license to produce several of the Intel processors.

In a licensing agreement, Intel shares with IBM the original mask, or design, of the chip. A CPU chip mask, which is the photographic blueprint of the processor, is used to etch the intricate signal pathways into a silicon chip. Using this mask, IBM can produce the chip as Intel designed it or make modifications in the basic design. Normally, IBM must also share with Intel any modifications that it makes.

Restrictions often exist in the way that IBM can market these Intel-derived chips. Other manufacturers (such as AMD, Harris, and Siemens) licensed several of the older Intel processors up through the 286, but no manufacturer other than IBM has licensed the 386 and newer processor masks. The 386 and newer chips produced by companies other than Intel and IBM are reverse-engineered without Intel's blessing. In fact, Intel has been involved in drawn-out legal battles with several of the clone-processor manufacturers.

Several years ago, IBM licensed the 386 processor mask. The company since has produced several variations that include modifications of the basic Intel design, including some versions that also carry a 486 designation. Some of the IBM processors have few design changes; others have major differences that make the chip faster, less power-hungry, or both. Table 6.12 summarizes the IBM processors that are derived from Intel masks.

IBM SLC Processors

The IBM SLC processors are enhanced versions of the Intel 386SX that offer greater performance and lower power consumption. The SLC processors, which are based on the Intel 386SX mask, have the physical form factor and pinout of the 386SX, with some of the unused pins now being used for cache control.

Maximum Memory	Integral Cache	Burst Mode	Integral FPU	No. of Transistors	Date Introduced
16M	8K	No	No	955,000	October 1991
16M	16K	No	No	1,349,000	1992
16M	16K	No	No	1,349,000	June 1992
4G	16K	No	No	1,400,000	1993
4G	16K	No	No	1,400,000	1993

The IBM-enhanced chips have some 486 features (including the internal cache) and perform as well as, or even better than, a 486SX at a much lower cost. IBM states that the 386SLC chip is as much as 80 percent faster than the standard 386SX chip, which means that a 20 MHz 386SLC can outperform a 33 MHz 386DX system. The first system to use any of these custom IBM processors was the PS/2 Model 57; these processors now are used in other manufacturers' systems as well.

The SLC processors are often criticized for having only a 16-bit data bus and a 24-bit address bus. The address bus allows access to only 16M of memory, which is fine for most users. Users who need more memory, however, should look to the IBM Blue Lightning (BL) or to the Intel 486 processors and beyond, which provide a full 32 bits of memory addressing.

Although the SLC processors have only a 16-bit data bus, these processors still outperform many 32-bit chips because of the large (8K or 16K) internal cache. In fact, the 16K cache in the 486SLC performed so well that they changed the designation of the chip to 486. This chip performs about as well as an Intel 486 chip running at equal clock rates, even though the IBM version is only 16 bits and the Intel version is 32 bits. Because all the SLC chips have the full 486 instruction set, including the cache-control instructions, the designation is accurate.

The SLC2 version runs at twice the system speed (clock-doubled) and yet costs less than the standard Intel 486 chips while providing greater performance. You can find the 486SLC2 processor in many bargain-priced motherboards that offer nearly the performance of the Intel 486SX2 and DX2 processors for much less money. Although the IBM processors do not have a built-in math coprocessor, as some of the Intel processors do, most of the motherboards sold with the IBM processors include an Intel or Intel-compatible 387SX math coprocessor.

Blue Lightning Processors

The IBM BL (Blue Lightning) processors are based on the 386DX mask and are full 32-bit processors. These processors also support a full 32-bit memory addressing scheme, which means that they can support the same 4G memory as the Intel 486 processors.

The IBM Blue Lightning was the first processor to run in clock-tripled mode, which is available in 25/75 MHz and 33/100 MHz versions. Notice that most 33 MHz systems actually run at 33.33 MHz, so a clock-tripled system effectively would run at 100 MHz.

All the IBM processors fully support power-management capabilities, and all of them are considered to be low-power (or green) processors. Due to their low power consumption and low-voltage design, these processors often are used in systems that meet the EPA Energy Star certification program, which sets standards for energy conservation and low power consumption in PC systems.

The SLC2 and BL processors are low-power units that run on 3.3 volts, due in part to IBM's superior chip-fabrication facilities. IBM reduced the die size and power consumption greatly by going to a 0.6-micron process for these chips. (This number refers to the minimum individual feature size on the chip mask. Intel did not use the a 0.6-micron process until the DX4 and Pentium chips were introduced.)

The original agreement with Intel was intended to prevent IBM from marketing its processors in the open market in competition with Intel. The agreement specifies that IBM cannot sell processors individually, but only as parts of modules. This restriction essentially means that the chip must be mounted on a circuit board of some type—which actually caused IBM to enter the motherboard and upgrade-board business.

Several companies have designed motherboards and processor upgrade boards that use the low-cost IBM processors. These boards are manufactured by IBM, but they often are designed by the companies for whom IBM makes them. Among the most popular of these motherboards are the Alaris motherboards, which use the IBM processors and an AMI BIOS; IBM manufactures these boards at its manufacturing plant in Charlotte, North Carolina.

It is important to note that for any of these products, IBM is acting as a subcontractor to the company that is considered to be the manufacturer; IBM is not responsible for the design, sales, or support of the products. Support must come from the (so-called) manufacturer, not from IBM.

Intel-Compatible Processors

Several companies—mainly AMD and Cyrix—have developed new processors that are compatible with one or more of the Intel processors. These chips are fully Intel-compatible, which means that they emulate every processor instruction in the Intel chips. Most of the chips are pin-compatible, which means that they can be used in any system designed to accept an Intel processor; others require a custom motherboard design. Any hardware or software that works on Intel-based PCs will work on PCs made with these third-party CPU chips.

AMD and Cyrix have developed their own versions of the Intel 386 and 486 processors in a variety of speeds and configurations.

Cyrix markets a chip called the 486DRx2, which, the company states, can be used as a direct replacement for a 386 chip in existing systems. Although the Cyrix chip lacks some of the important features of the Intel 486, it gives a 386-based system many of the benefits of the 486.

Math Coprocessors

The next several sections cover the math coprocessor. Each central processing unit designed by Intel (and cloned by other companies) can use a math-coprocessor chip, although the Pentium and 486 chips have a built-in math coprocessor. Coprocessors provide hardware for floating-point math, which otherwise would create an excessive drain on the main CPU. Math chips speed your computer's operation only when you are running software that is designed to take advantage of the coprocessor.

Math chips (as coprocessors sometimes are called) can perform high-level mathematical operations—long division, trigonometric functions, roots, and logarithms, for example—at 10 to 100 times the speed of the corresponding main processor. Math chips also are more accurate in these calculations than are the integer-math units built into the primary CPU. The integer units in the primary CPU work with real numbers, so they perform addition, subtraction, and multiplication operations. The primary CPU is designed to handle such computations; these operations are not offloaded to the math chip.

The instruction set of the math chip is different from that of the primary CPU. A program must detect the existence of the coprocessor and then execute instructions written explicitly for that coprocessor; otherwise, the math coprocessor draws power and does nothing else. Fortunately, most modern programs that can benefit from the use of the coprocessor correctly detect and use the coprocessor. These programs usually are math-intensive programs: spreadsheet programs; database applications; statistical programs; and some graphics programs, such as computer-aided design (CAD) software. Word processing programs
do not benefit from a math chip and therefore are not designed to use one.

Table 6.13 summarizes the coprocessors available for the Intel family of processors.

Table 6.13 Math Coprocessor Summary

Processor	Coprocessor
8086	8087
8088	8087
286	287
386SX	387SX
386SL	387SX
386SLC	387SX
486SLC	387SX
486SLC2	387SX

386DX	387DX
486SX	487SX, DX2/Overdrive*
487SX*	Built-in FPU
486SX2	DX2/Overdrive**
486DX	Built-in FPU
486DX2	Built-in FPU
486DX4	Built-in FPU
Pentium	Built-in FPU

FPU = Floating-Point Unit
**The 487SX chip is a modified pinout 486DX chip with the math coprocessor enabled. When you plug in a 487SX chip, it disables the 486SX main processor and takes over all processing. This chip has been discontinued and is replaced by the DX2/Overdrive processor.*
***The DX2/Overdrive is equivalent to the SX2 with the addition of a functional FPU.*

Within each 8087 group, the maximum speed of the math chips vary. A suffix digit after the main number, as shown in table 6.14, indicates the maximum speed at which a system can run a math chip.

Table 6.14 Maximum Math-Chip Speeds

Part	Speed
8087	5 MHz
8087-3	5 MHz
8087-2	8 MHz
8087-1	10 MHz
80287	6 MHz
80287-6	6 MHz
80287-8	8 MHz
80287-10	10 MHz

The 387 math coprocessors, as well as the 486 or 487 and Pentium processors, always indicate their maximum speed rating in MHz in the part-number suffix. A 486DX2-66, for example, is rated to run at 66 MHz. Some processors incorporate clock multiplication, which means that they may run at different speeds compared with the rest of the system.

The performance increase in programs that use the math chip can be dramatic—usually, a geometric increase in speed occurs. If the primary applications that you use can take advantage of a math coprocessor, you should upgrade your system to include one.

Most systems that use the 386 or earlier processors are socketed for a math coprocessor as an option, but they do not include a coprocessor as standard equipment. A few systems on the market don't even have a socket for the coprocessor because of cost and size considerations. Usually, these systems are low-cost or portable systems, such as older laptops, the IBM PS/1, and the PCjr. For more specific information about math coprocessors, see the discussions of the specific chips—8087, 287, 387, and 487SX—in the following sections.

Table 6.15 shows some of the specifications of the various math coprocessors.

Table 6.15 Intel Math Coprocessor Specifications

Name	Power Consumption	Case Min. Temp.	Case Max. Temp.	No. of Transistors	Date Introduced
8087	3 watts	0°C, 32°F	85°C, 185°F	45,000	1980
287	3 watts	0°C, 32°F	85°C, 185°F	45,000	1982
287XL	1.5 watts	0°C, 32°F	85°C, 185°F	40,000	1990
387SX	1.5 watts	0°C, 32°F	85°C, 185°F	120,000	1988
387DX	1.5 watts	0°C, 32°F	85°C, 185°F	120,000	1987

Most often, you can learn what CPU and math coprocessor are installed in a particular system by checking the system documentation. The following section examines the Intel family of CPUs and math coprocessors in more detail.

8087 Coprocessor

Intel introduced the 8086 processor in 1976. The math coprocessor that was paired with the chip—the 8087—often was called the numeric data processor (NDP), the math coprocessor, or simply the math chip. The 8087 is designed to perform high-level math operations at many times the speed and accuracy of the main processor. The primary advantage of using this chip is the increased execution speed in number-crunching programs, such as spreadsheet applications. Using the 8087 has several minor disadvantages, however, including software support, cost, power consumption, and heat production.

The primary disadvantage of installing the 8087 chip is that you notice an increase in speed only in programs that are written to use this coprocessor—and then not in all operations. Only math-intensive programs such as spreadsheet programs, statistical programs, CAD software, and engineering software support the chip. Even then, the effects vary from application to application, and support is limited to specific areas. For example, versions of Lotus 1-2-3 that support the coprocessor do not use the coprocessor for common operations such as addition, subtraction, multiplication, and division.

Applications that usually do not use the 8087 at all include word processing programs, telecommunications software, database programs, and presentation-graphics programs.

To test the speed capabilities of the 8087 math coprocessor, two spreadsheets were created, each with 8,000 cells. The first spreadsheet used simple math tasks—addition, subtraction, multiplication, and division—split evenly among the 8,000 cells. The second spreadsheet used high-level math operations—including formulas that used SQRT, SIN, COS, and TAN calculations—throughout the 8,000 cells. The following table shows the recalculation times.

Spreadsheet	XT without 8087	XT with 8087
Sheet 1 (standard math)	21 seconds	21 seconds
Sheet 2 (high-level math)	195 seconds	21 seconds

The addition of an 8087 to a standard IBM XT did nothing for the spreadsheet that contained only simple math, but the 8087 calculated the spreadsheet that contained high-level math in one-tenth the time. (This also was the time it took both systems to calculate the spreadsheet that contained only simple math.)

If your spreadsheets consist of nothing but addition, subtraction, multiplication, and division calculations, before buying a math chip, you need to know whether the program that you use takes advantage of the math chip for simple calculations. Many new programs are designed to support the 8087 chip for these operations. Installing an 8087 can extend the useful life of a PC or XT, because the chip closes some of the performance gaps between PC- or XT-type computers and AT-type computers. In short, the chip is an asset whenever the software supports it.

The 8087 chip is inexpensive, costing as little as $50. Remember to purchase a chip with the correct maximum-speed rating; the 8087 must be rated to run at the same rate of speed as the CPU or faster, because the main CPU and the coprocessor must run in synchronization. In an IBM XT, for example, the 8088 and the 8087 run at 4.77 MHz. Look in your system documentation to find the speed at which your system will run the math chip.

Math chips are quite power-hungry because of the number of transistors included. A typical 8088 has only about 29,000 transistors, but the 8087 has about 45,000. (Nearly all the 45,000 transistors are dedicated to math functions, which is why the 8087 can perform math operations so well.) This figure translates to nearly double the calculating horsepower, as well as double the electrical power drain. In a heavily loaded PC, the 8087 could be the straw that breaks the camel's back: the power supply might be insufficient for the increased load. The chip draws nearly one-half amp of current.

Another problem is the amount of heat generated: 3 watts. A 75 MHz 486DX4 with 1.6 million transistors uses only 3.63 watts. This heat level generated in such a small chip can raise the temperature to more than 180° Fahrenheit. (The approved maximum temperature for most 8087s is 185°F.) For this reason, math coprocessor chips usually are made of ceramic.

Power and heat are not problems in XT or portable systems, because these systems are built to handle these situations. The PC, however, usually requires a higher-watt power supply with a more powerful cooling fan to handle the load. Power supplies are covered later in this chapter.

80287 Coprocessor

Imagine that two company employees have computers. One user has an IBM XT, and the other has a 6 MHz AT. Both employees use Lotus 1-2-3 as their primary application. The AT user delights in being able to outcalculate the XT user by a factor of three. The XT user purchases an 8087 math chip for $50 and installs it. The XT user then finds that the XT calculates many spreadsheets 10 times faster than before—or more than three times as fast as the AT.

This performance frustrates the AT user, who thought that the AT was the faster system. The AT user, therefore, purchases an 80287 for $50 and discovers that the AT is merely equal in speed to the XT for many spreadsheet recalculations. In a few situations, however, the XT may still outrun the AT.

Of course, the AT user wants to know why the 80287 chip did not make the AT superior to the XT, by a significant margin, for spreadsheet recalculations. (For "normal" processing, which does not use the math chip's high-level math functions, the AT holds its performance edge.) The answer is in the 80287 chip. For various design reasons, the 8087 chip has much more effect on the speed of the PC and XT than the 80287 has on the AT.

The 80287, internally, is the same math chip as the 8087, although the pins used to plug them into the motherboard are different. Because the AT has a healthy power supply and generous, thermostatically controlled fan cooling, the heat and power problems mentioned in the discussion of the 8087 generally don't apply to the 287. Internally, however, the 80287 and the 8087 operate as though they were identical.

Another reason is that the 80286 and its math chip are *asynchronous*, which means that the chips run at different speeds. The 80287 math chip usually runs at two-thirds the speed of the CPU. In most systems, the 80286 internally divides the system clock by 2 to derive the processor clock. The 80287 internally divides the system-clock frequency by 3. For this reason, most AT-type computers run the 80287 at one-third the system clock rate, which also is two-thirds the clock speed of the 80286. Because the 286 and 287 chips are asynchronous, the interface between the 286 and 287 chips is not as efficient as with the 8088 and 8087.

In summary, the 80287 and the 8087 chips perform about the same at equal clock rates. The original 80287 is not better than the 8087 in any real way—unlike the 80286, which is superior to the 8086 and 8088. In most AT systems, the performance gain that you realize by adding the coprocessor is much less substantial than the same type of upgrade for PC- or XT-type systems or for the 80386.

Some systems run the 80287 and the 8087 at the same speed. PS/2 Models 50, 50 Z, and 60 use circuitry that enables both the 80286 and 80287 to run at 10 MHz. The PS/2 Model 25-286 and 30-286, however, follow the standard AT-type design, in which the 286 runs at 10 MHz and the 287 runs at 6.67 MHz.

You must consult the system documentation to learn the speeds at which your system would run a 287 coprocessor, because the motherboard designers determine these specifications. Table 6.16 shows 80286 and 80287 clock speeds for most AT-type systems.

Table 6.16 80286 and 80287 Clock Speeds (in MHz)

System Clock	80286 Clock	80287 Clock
12.00	6.00	4.00
16.00	8.00	5.33
20.00	10.00	6.67
24.00	12.00	8.00
32.00	16.00	10.67

How can you improve this differential in performance gain? One method takes advantage of the fact that the 80286 and the 80287 run asynchronously. You can install an add-in board that uses its own clock signal to drive the 80287 chip and therefore can drive the chip at any speed. Some companies have designed a simple speed-up circuit that includes
a crystal and an 8284 clock generator chip, all mounted on special boards, some of which are not much bigger than the 287 socket. This special board, called a *daughterboard*, is plugged into the 287 socket; then the 287 is plugged in on top of the special board. Because such boards separate the crystal and clock generator from the motherboard circuitry, the daughterboard can run the 80287 at any speed you want, up to the maximum rating of
the chip—8 MHz, 10 MHz, 12 MHz, or more—without affecting the rest of the system.

You could, for example, add one of these daughterboards to your old 6 MHz AT and run a 287 at 10 MHz. Without the daughterboard, the chip would run at only 4 MHz. The boards are available from many math-coprocessor suppliers. Using these boards is highly recommended if you run math-intensive programs. Remember that this type of speed increase does not apply to systems that use 8087 or 80387 chips, because these systems must run the math coprocessor at the same speed as the main CPU.

Intel has introduced new variations of the 80287 called the 287XL and 287XLT. (The original 287 has been discontinued; only the 287XL and XLT are available today.)
The XL version is designed to be a replacement for the standard 287 math coprocessor. The XLT version is functionally identical to the XL but has a plastic leadless chip carrier (PLCC) case, which some laptop systems require.

These redesigned XL chips are patterned after the 387 instead of the 8087. The XL chips consume much less power than the original 287 chips, because they are constructed with CMOS technology. The XL chips perform about 20 percent faster than the original 287 at any clock rate as a result of their improved design. The design improvements extend to the instruction set, which includes 387 trigonometric functions that are not available with older 287 coprocessors.

The XL chips are available at only one speed rating: 12.5 MHz. The chips can be run at lower speeds on slower systems. Unlike the 287 daughterboards, these chips do not increase the clock speed of a math chip.

Many older diagnostics programs incorrectly identify the XL chips because they were designed after the 387 math chips. Some diagnostics simply indicate that the 287XL is a 387; other diagnostics incorrectly show a problem with the math coprocessor if a 287XL is installed. Intel provides a special diagnostics program called CHKCOP (which stands for CHecK COProcessor) that can test all its math coprocessors. You can get this program on disk from Intel's customer-support department or download it from the Intel BBS at (503) 645-6275.

After considering all these issues, if you decide to invest in a 287 chip, remember that only the XL or XLT versions are available now and that they are rated for up to 12.5 MHz operation. Adding the 287 to an AT is a good idea if the software that you use supports the chip. You also should consider using one of the math coprocessor-speedup daughterboards, which will run the newer XL chips at the maximum 12 MHz rating regardless of your system's clock speed. Otherwise, the benefits may not be enough to justify the cost.

80387 Coprocessor

Although the 80387 chips run asynchronously, 386 systems are designed so that the math chip runs at the same clock speed as the main CPU. Unlike the 80287 coprocessor, which was merely an 8087 with different pins to plug into the AT motherboard, the 80387 coprocessor is a high-performance math chip designed specifically to work with the 386.

All 387 chips use a low-power-consumption CMOS design. The 387 coprocessor has two basic designs: the 387DX coprocessor, which is designed to work with the 386DX processor; and the 387SX coprocessor, which is designed to work with the 386SX, SL, or SLC processors.

Originally, Intel offered several speeds for the 387DX coprocessor. But when the company designed the 33 MHz version, a smaller mask was required to reduce the lengths of the signal pathways in the chip. (A *mask* is the photographic blueprint of the processor and is used to etch the intricate signal pathways into a silicon chip.) Intel reduced the feature size from 1.5 to 1 micron. This action reduced the size of the silicon chip by 50 percent.

In addition to the size reduction, other design improvements were engineered into the new mask, resulting in a 20-percent improvement in processing efficiency. The 33 MHz version, therefore, outperformed other versions even at slower clock rates.

At the time, purchasing the 33 MHz version of the 387DX was a good idea (even for a 20 MHz 386 system), because the chip would run 20 percent faster than a 20 MHz 387. In October 1990, however, Intel upgraded the entire 387DX line to the improved mask, resulting in a 20 percent performance boost across the board.

You can easily identify these improved 387DX coprocessors by looking at the 10-digit code below the 387 part number. The older (slower) chips begin this line with the letter *S*, and the improved (faster) chips do not. Recently, Intel discontinued all its 387DX processors except the 33 MHz version, which, of course, always used the new design. (Remember that even though the chip is rated for 33 MHz, it runs at any lower speed.)

The 387SX coprocessors are designed to work specifically with 386SX, SL, or SLC processors. All versions of the 387SX use the improved mask design. When you select a 387SX for your system, be sure to purchase one that is rated at a speed equal to or higher than that of your CPU. Currently, Intel 387SX chips are available at speeds up to 25 MHz.

Note

Because Intel lagged in developing the 387 coprocessor, some early 386 systems were designed with a socket for a 287 coprocessor. Performance levels associated with that union, however, leave much to be desired.

Installing a 387DX is easy, but you must be careful to orient the chip in its socket properly; otherwise, the chip will be destroyed. The most common cause of burned pins on the 387DX is incorrect installation. In many systems, the 387DX is oriented differently from other large chips. Follow the manufacturer's installation instructions carefully to avoid damaging the 387DX; Intel's warranty does not cover chips that are installed incorrectly.

Several manufacturers have developed their own versions of the Intel 387 coprocessors, some of which are touted as being faster than the original Intel chips. The general compatibility record of these chips is very good. Intel has significantly reduced the prices of its own coprocessors, however, which means that these third-party chips usually are only a few dollars cheaper than the Intel version.

When the 387s were introduced, the Intel 33 MHz 387DX chips listed for more than $2,000. Today, you can buy the chip from various suppliers for as little as $90. The cost is so low that many people should consider making this upgrade. If the software that you run supports the chip, the performance gains can be very impressive.

Weitek Coprocessors

In 1981, several Intel engineers formed Weitek Corporation. Weitek has developed math coprocessors for several systems, including those based on Motorola processor designs. Intel originally contracted Weitek to develop a math coprocessor for the Intel 386 CPU, because Intel was behind in its own development of the 387 math coprocessor. The result was the Weitek 1167, a custom math coprocessor that uses a proprietary Weitek instruction set that is incompatible with the Intel 387.

The Weitek 1167 is not a single chip; it is a daughterboard consisting of several chip elements that plug into a special 112-pin Weitek socket. To use the Weitek processors, your system must have the required socket, which is incompatible with the 387 math coprocessor and 486SX processor enhancement sockets. The daughterboard includes a socket for an Intel 387 coprocessor so that both coprocessors can be installed in a system; as a result, software that runs either Weitek or Intel math instructions will work on such a system.

The 1167 was replaced in April 1988 by a single-chip version called the 3167. Many computers, such as the Compaq 386, contain a special socket that enables you to use a Weitek 3167 math coprocessor or an Intel 387DX. This socket has three rows of holes on all four sides. The inner two rows of pins are compatible with the Intel 387DX. If you want to install a 387DX in the special socket, however, you must use extreme caution to orient the chip correctly; otherwise, you could damage both the computer and the 387DX.

Read your system documentation to determine the correct procedure for installing the 387DX in your computer. Some computers, such as the Tandy 4000, use the Weitek socket but do not support the 387DX. Contact your computer manufacturer or dealer for more specific information.

Unfortunately, even if you have the socket for the Weitek processor, your software probably does not support it. As mentioned, your software must contain programming code that makes use of the specific capabilities of a math coprocessor.

Weitek introduced the 4167 coprocessor chip for 486 systems in November 1989. To use the Weitek coprocessor, your system must have the required additional socket. Before purchasing one of the Weitek coprocessors, you should determine whether your software supports it; then you should contact the software company to determine whether the Weitek has a performance advantage over the Intel coprocessor.

80487 Upgrade

The Intel 80486 processor was introduced in late 1989, and systems using this chip appeared during 1990. The 486DX integrated the math coprocessor into the chip.

The 486SX began life as a full-fledged 486DX chip, but Intel actually disabled the built-in math coprocessor before shipping the chip. As part of this marketing scheme, Intel marketed what it called a 487SX math coprocessor. Motherboard manufacturers installed an Intel-designed socket for this so-called 487 chip. In reality, however, the 487SX math chip was a special 486DX chip with the math coprocessor enabled. When you plugged this chip into your motherboard, it disabled the 486SX chip and gave you the functional equivalent of a full-fledged 486DX system.

Perhaps that somewhat strange marketing scheme is responsible for some of the confusion caused by the Intel advertisements that feature a 486SX system with a neon vacancy sign pointing to an empty socket next to the CPU chip. Unfortunately, these ads do not transmit the message properly. Few people seem to understand that the socket next to the CPU in a 486SX system is not a math-coprocessor socket, but an Overdrive socket.

Essentially, systems that have this extra socket have two processor sockets; however, you can use only one of them at any time. When you install a chip in the secondary socket, that chip takes over from the primary processor and puts the original processor to sleep.

Most newer 486SX systems use a surface-mounted PQFP (Plastic Quad Flat Pack) or SQFP (Small Quad Flat Pack) version that is permanently soldered to the motherboard. These systems still have a conventional processor socket of a specific design for new Overdrive processors. These Overdrive chips will contain all processing functions, including the math-coprocessor functions, and will shut down the 486SX when they are installed. Depending on the type of processor socket on the motherboard, you can install a DX2 or DX4 processor, or even a special version of the Pentium chip.

For more information on this subject, see the section on Overdrive processors earlier in this chapter.

Processor Tests

The processor is easily the most expensive chip in the system. Processor manufacturers use specialized equipment to test their own processors, but you have to settle for a little less. The best processor-testing device to which you have access is a system that you know is functional; you then can use the diagnostics available from IBM and other system manufacturers to test the motherboard and processor functions. Most systems mount processors in a socket for easy replacement.

Landmark offers specialized diagnostic software called Service Diagnostics to test various processors. Special versions are available for each processor in the Intel family. If you don't want to purchase this kind of software, you can perform a quick-and-dirty processor evaluation by using the normal diagnostic program supplied with your system.

Because the processor is the brain of a system, most systems don't function with a defective one. If a system seems to have a dead motherboard, try replacing the processor with one from a functioning motherboard that uses the same CPU chip. You may find that the processor in the original board is the culprit. If the system continues to play dead, however, the problem is elsewhere.

Known-Defective Chips

A few system problems are built in at the factory, although these bugs or design defects are rare. By learning to recognize one of these problems, you may avoid unnecessary repairs or replacements. This section describes several known defects in system processors.

Early 8088s. A bug in some early 8088 processors allowed interrupts to occur after a program changed the stack segment register. (An interrupt usually is not allowed until the instruction after the one that changes the stack segment register.) This subtle bug may cause problems in older systems. Most programmers have adopted coding procedures that work around the bug, but you have no guarantee that these procedures exist in all software programs.

Another problem is that the bug may affect the operation of an 8087 math coprocessor. Approximately 200,000 IBM PC units sold during 1981 and 1982 were manufactured with the defective chip.

Originally, in the 8087 math-coprocessor-chip package, IBM always included an 8088 to be installed with the math chip. This practice led to rumors that the parts somehow matched. The rumors were unfounded; IBM had simply found an easy way to prevent machines that used its 8087 chips from using the defective 8088. Because the cost of the chip was negligible, IBM included a bug-free 8088 and eliminated many potential service problems.

You can check the 8088 chip with diagnostic software, or you can identify a good or bad chip by its appearance. If you can open the unit to look at the 8088 chip, the manufacturer and copyright date printed on the chip provide clues about which version you have. An 8088 chip made by a manufacturer other than Intel is bug-free, because Intel began licensing the chip mask to other manufacturers after the bug was corrected. If a chip was manufactured by Intel, older (defective) parts have a 1978 copyright date; newer (good) parts have 1978 and 1981 (or later) copyright dates.

This marking on Intel 8088 chips indicates a chip that has the interrupt bug:

8088

(c)INTEL 1978

The following markings on Intel 8088 chips indicate chips on which the bug is corrected:

8088

(c)INTEL '78 '81

8088

(c)INTEL '78 '83

Many diagnostic programs can identify the chip. You also can identify the chip yourself by using DEBUG, which comes in DOS versions 2.0 and later. Just load DEBUG at the prompt, and enter the commands shown in the following example. The commands that you enter appear in boldface type; the DEBUG screen output is not. XXXX indicates a segment address, which varies from system to system.

A 100

[XXXX:0100] MOV ES,AX

[XXXX:0102] INC AX

[XXXX:0103] NOP

[XXXX:0104

T

AX=0001 BX=0000 CX=0000 DX=0000 SP=FFEE BP=0000 SI=0000 DI=0000

DS=XXXX ES=0000 SS=XXXX CS=XXXX IP=0103 NV UP EI PL NZ NA PO NC

XXXX:0103 90 NOP

—Q

The A 100 command tells DEBUG to assemble some instructions, of which three are entered. The T command then executes a Trace, which normally executes a single instruction, displays the contents of the 8088's registers, and then stops. The Trace command usually executes only one instruction. However, when the instruction is MOV to a segment register, as in this case, the Trace command should execute the second instruction before interrupting the program. The third instruction is a dummy no-operation instruction.

Look at the value shown by DEBUG for the register AX. If AX is equal to 0000, the processor has a bug. If AX is 0001, the second instruction in the test was executed properly, and the chip is good. If the second instruction is executed, DEBUG increments the value of the AX register by 1. In this example, after the Trace is executed, AX equals 0001, which indicates a good chip.

Note

If you try this test on a 286 or higher system, the test will fail; it is valid only for 8088s.

If you have an 8087 and a 4.77 MHz 8088 dated '78 or an 8088 that fails this test, you can get a free replacement 8088. Contact Intel customer support for the replacement. Only 4.77 MHz 8088 chips may need to be upgraded. The 8088-2 and 8088-1 do not require replacement. You also can purchase a replacement 8088 from most chip houses for less than $10. If you suspect that your chip is bad, a replacement is inexpensive insurance.

Early 80386s. Some early 16 MHz Intel 386DX processors had a small bug that you may encounter in troubleshooting what seems to be a software problem. The bug, which apparently is in the chip's 32-bit multiply routine, manifests itself only when you run true 32-bit code in a program such as OS/2 2.x, UNIX/386, or Windows in Enhanced mode. Some specialized 386 memory-management software systems also may invoke this subtle bug, but 16-bit operating systems (such as DOS and OS/2 1.x) probably will not.

The bug usually causes the system to lock up. Diagnosing this problem can be difficult, because the problem generally is intermittent and software-related. Running tests to find the bug is difficult; only Intel, with proper test equipment, can determine whether your chip has a bug. Some programs can diagnose the problem and identify a defective chip, but they cannot identify all defective chips. If a program indicates a bad chip, you certainly have a defective one; if the program passes the chip, you still may have a defective one.

Intel requested that its 386 customers return possibly defective chips for screening, but many vendors did not return them. Intel tested returned chips and replaced defective ones. The known-defective chips later were sold to bargain liquidators or systems houses that wanted chips that would not run 32-bit code. The known-defective chips were stamped with a 16-bit SW Only logo, indicating that they were authorized to run only

16-bit software.

Chips that passed the test, and all subsequent chips produced as bug-free, were marked with a double-sigma ( ) code, which indicates a good chip. 386DX chips that are not marked with either 16-bit SW Only or the designation have not been tested by Intel and may be defective.

The following marking indicates that a chip has not yet been screened for the defect; it may be either good or bad. Return a chip of this kind to the system manufacturer, which will return a free replacement.

80386-16

The following marking indicates that the chip has been tested and has the 32-bit multiply bug. The chip works with 16-bit software (such as DOS) but not with 32-bit, 386-specific software (such as Windows or OS/2).

80386-16

16-bit SW Only

The following mark on a chip indicates that it has been tested as defect-free. This chip fulfills all the capabilities promised for the 80386.

80386-16

This problem was discovered and corrected before Intel officially added DX to the part number. So if you have a chip labeled as 80386DX or 386DX, it does not have this problem.

Another problem with the 386DX can be stated more specifically. When 386-based versions of XENIX or other UNIX implementations are run on a computer that contains a 387DX math coprocessor, the computer locks up under certain conditions. The problem does not occur in the DOS environment, however. For the lockup to occur, all the following conditions must be in effect:

- Demand page virtual memory must be active.
- A 387DX must be installed and in use.
- DMA (direct memory access) must occur.
- The 386 must be in a wait state.

When all these conditions are true at the same instant, the 386DX ends up waiting for the 387DX, and vice versa. Both processors will continue to wait for each other indefinitely. The problem is in certain versions of the 386DX, not in the 387DX math coprocessor.

Intel published this problem (Errata 21) immediately after it was discovered, to inform its OEM customers. At that point, it became the responsibility of each manufacturer to implement a fix in its hardware or software product. Some manufacturers, such as Compaq and IBM, responded by modifying their motherboards to prevent these lockups from occurring.

The Errata 21 problem occurs only in the B Stepping version of the 386DX and not in the later D Stepping version. You can identify the D Stepping version of the 386DX by the letters DX in the part number (for example, 386DX-20). If DX is part of the chip's part number, the chip does not have this problem.

Other Processor Problems

Some other problems with processors and math coprocessors are worth noting.

After you remove a math coprocessor from an AT-type system, you must rerun your computer's SETUP program. Some AT-compatible SETUP programs do not properly unset the math-coprocessor bit. If you receive a Power-On Self Test (POST) error message because the computer cannot find the math chip, you may have to unplug the battery from the system board temporarily. All SETUP information will be lost, so be sure to write down the hard drive type, floppy drive type, and memory and video configurations before unplugging the battery. This information is critical in reconfiguring your computer correctly.

Another strange problem occurs in some IBM PS/2 Model 80 systems when a 387DX is installed. In the following computers, you may hear crackling or beeping noises from the speaker while the computer is running:

- 8580 Model 111, with serial numbers below 6019000
- 8850 Model 311, with serial numbers below 6502022

If you are experiencing this problem, contact IBM for a motherboard replacement.

Heat and Cooling Problems

Heat can be a problem in any high-performance 486 or Pentium system. The Intel DX2/66 consumes about 40 percent more power than a 33 MHz 486DX and generates correspondingly more heat. The Pentium generates even more heat. If your system is based on the DX2 or Pentium chip, you must dissipate the extra thermal energy; the fan inside your computer case may not be able to handle the load.

To cool a system in which processor heat is a problem, you can buy (for less than $20) a special attachment for the CPU chip called a heat sink, which draws heat away from the CPU chip. Many applications may need only a larger standard heat sink with additional or longer fins for a larger cooling area. Several heat-sink manufacturers are listed in Appendix B, the "Vendor List."

I prefer clip-on heat sinks, but some are attached with a special adhesive. In many cases, you need to use a thermal transfer paste even if the heat sink is clipped on. This paste fills any small air gaps between the processor and the heat sink, providing a more effective transfer of heat.

Most of the Overdrive processors that Intel will be introducing over the next few years will have a built-in active heat sink, which includes a fan. Unlike the aftermarket add-on