Intel Describes P5 Internal Architecture

Cache Details, Floating Point Timing, Reliability Features Revealed

By Linley Gwennap

Continuing its six-month striptease, Intel revealed several new details about its upcoming P5 chip (now officially named "Pentium") during its presentation at the Microprocessor Forum. Although much of this information was already available *(see **061201.PDF**)*, this session provided the first complete description of the internal architecture of the P5, Intel's follow-on to the 486. Highlights of the presentation, made by Intel fellow John Crawford, included details on the cache configuration, floating-point timing, and new reliability features.

Figure 1 shows Intel's block diagram of the CPU. For the first time, Intel revealed that the separate instruction and data caches are both 8K in size and that the data cache is eight-way interleaved. Instead of using additional die area to increase the size of the caches, Intel created an unusual dual-access data cache by dual-porting the cache tag arrays and TLBs. Thus, the cache can provide two 32-bit values each cycle as long as there are no bank conflicts. The high degree of interleaving greatly reduces the number of conflicts. The instruction cache is single-access but can supply 256 bits (32 bytes) to the prefetch buffer in a single cycle, twice as many as the 486's unified cache.

Crawford also confirmed some details about the P5 pipeline. The dual integer units allow two "simple" instructions to be issued each clock cycle, including register-register and register-memory instructions. Crawford said that some instructions require multiple cycles in the "execute" phase; in this situation, the second pipeline is also frozen, preventing another instruction from being issued. He would not confirm if register-memory instructions take two cycles, as they do on the 486.

He also gave the latencies for floating-point operations: 3 cycles for add and multiply, and 18–38 cycles for divide (depending on the precision). The floating-point unit supports 80-bit extended precision.

Intel also revealed an impressive array of error detection features in the P5. Although the 486 provides parity checking for the external data bus, the P5 adds parity for the external address bus as well. In addition, all internal P5 data structures are parity-protected, including the code and data caches, cache tags, TLBs, microcode ROM, and the branch target buffer. For ultimate security, fault-tolerant systems can configure one P5 chip to check the outputs of another P5 on a cycle-by-cycle basis, signaling any inconsistency.

To unleash the full performance of the P5, recompilation will be necessary. On integer code, Crawford said that code compiled with the "best" 486 compiler will run about 15% slower than code compiled for the P5, and "typical" x86 code will run about 30% slower. Recompiling is even more important for floating-point applications to take advantage of parallel FXCH execution; old binaries will run at just half the speed of recompiled code on the P5. Even without recompiling, however, users will still see a speedup over a 486.

An important question is whether programs that are recompiled for the P5 will still run efficiently on other x86 systems. Intel's figures show that integer performance is about the same on a 486 with either the P5 compiler or the best 486 compiler. Floating-point performance is about 10% higher with the P5 compiler on a 486 processor. It is likely that some of the P5 optimizations (such as relying on "simple" instructions) are detrimental to 486 performance, but the overall impact of all optimizations is about even.

Crawford's presentation did not discuss the bus interface at all, other than to confirm that it is a new, 64-bit bus. He revealed no new information about the branch prediction unit. Nothing was said regarding instruction set extensions; these will be kept under wraps as long as possible to make life difficult for other x86 chip vendors. Informa-

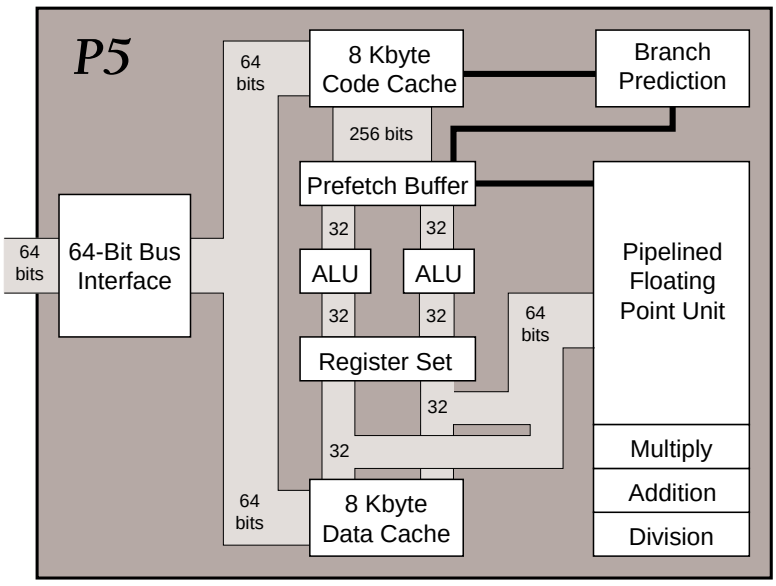

Figure 1. P5 internal block diagram.

tion on performance, pricing, and availability will probably be left until the official product announcement. Intel remains committed to a 1Q93 introduction.

Despite Crawford's reluctance, industry sources have revealed a few more tidbits about the P5. Taking a page from current RISC CPUs, the P5 will be able to map 2M chunks of memory with a single page-table entry. This feature has proven useful for mapping large operating systems and graphics frame buffers. The chip will also include a new 36-bit linear-addressing mode. Other high-performance CPUs use a similar scheme to support more than 4G of physical memory. Finally, Intel has also improved interrupt handling in virtual-8086 mode to reduce overhead; this should speed up applications running in DOS emulation mode under Windows or Windows NT.

P5 Target Markets

The feature set of the P5 indicates that it is poised to spring on a variety of markets. The largest, of course, is the x86 upgrade market. The typical PC user, who does not buy new P5-compiled software and uses very little floating point, will see a 40% to 80% performance increase over a 66-MHz 486DX2, and a bigger improvement over other 486 systems. Assuming Intel's standard pricing policies, this performance boost will be enough to migrate customers to the P5 over time.

The P5 will also provide a new weapon against other x86 vendors. Intel used the 286 ads to move customers to the 386, and the company is now wielding low 486 prices to incite them to move away from 386 chips. As AMD and others ready their versions of the 486, Intel will use the P5 to stay one step ahead of the pack.

The P5's dramatic increase in floating-point performance sends a message to RISC vendors that they can't take the workstation market for granted. The workstation market is too small to be very interesting to Intel, but Intel doesn't want SPARC or MIPS chips to take away personal computer sales, as the ACE group wants to do. Just as Bill Clinton campaigns in Texas to force George Bush to pay attention to his home turf, the P5 will force Sun and SGI to spend resources on upgrading the floating-point performance of their chips to avoid losing workstation sales to P5-based systems.

The error detection features and large physical memory support make the P5 more attractive to vendors such as NCR (AT&T), Sequent, and Unisys that build large, commercial systems using x86 processors. Although this market is also quite small, these companies also sell (and use) large numbers of PCs. Vendors see an advantage in using the same processor from the top of the product line to the bottom, so adding reliability features to the P5 protects the entire product line for Intel.

One area that Intel has not addressed at this point is improved graphics performance. Although the company admits the importance of graphics by including it as a significant component of the new iCOMP rating *(see 061302.PDF)*, Intel has not revealed any P5 features that are specific to graphics. Increased CPU power will, of course, improve graphics performance somewhat, and moving to a 64-bit bus will improve bandwidth. But the iCOMP formula hints at future disclosures, perhaps on-chip BitBLT support.

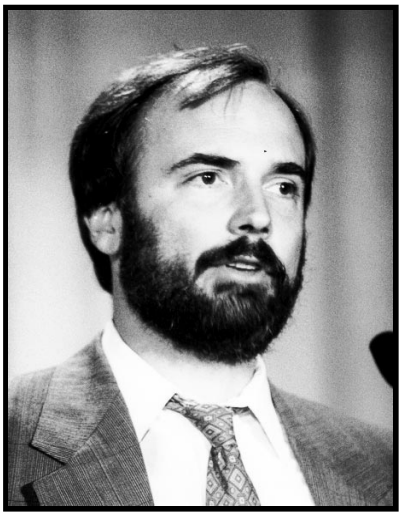

PHOTO BY ANNE HAMERSKY

"The demise of the x86 product line is greatly exaggerated.... There are a lot of techniques that are applicable not only to RISC architectures but also to ours, and there's a lot of speedups available, a lot of good techniques yet to be used.... We see no end in sight to improving this performance and maintaining our performance ramp of doubling every 18 months."

John Crawford, Intel

Conclusions

The most interesting feature of the P5 is its dual-access data cache. Because x86 code generates a relatively high number of data references per instruction, other x86 processors will also have to adopt this technique as they try to reach P5 levels of performance. Even RISC chips will need to improve memory bandwidth as they increase the number of instructions that they can execute in a cycle. Intel believes that the dual-access cache provides better performance than a larger, single-access cache using the same die area.

The new reliability features are overkill from a technical standpoint. The traditional belief is that microprocessors don't fail. With over 3 million transistors on the chip, the P5 team was probably wise to protect the large caches, but is parity for the microcode ROM really needed? Recently, however, customers have been demanding parity and even ECC protection as IBM has emphasized these features on its systems, so the moves in this area were probably dictated by marketing.

From a business perspective, the P5 will help protect Intel's installed base from clone chips while attacking the RISC chips' dominance of the workstation and business server markets. One key factor remains unannounced—the price. Intel will probably begin the P5 at a relatively high price, maximizing its profit margins. If Intel feels threatened, however, it can use an aggressively-priced P5 to pummel the forthcoming 486 competitors. ♦

