AMD Has Super Plans for Socket 7

K6 to Gain 100-MHz Bus, FP MMX, and On-Chip L2 Cache

by Brian Case

Making good on its promise to boost the performance of Socket 7, AMD revealed its 1998 roadmap at the recent Microprocessor Forum. The company's plans for the K6, outlined by Greg Favor, include a 100-MHz external processor bus, faster internal clock rates, superscalar MMX execution, a new MMX-like instruction set for floating-point operands, and a large on-chip level-2 cache. In addition, CEO Jerry Sanders dropped a few tidbits about the K7.

In 1H98, AMD will introduce the first of two new K6 chips, the K6 3D. This chip will implement the 100-MHz external bus and AMD's own floating-point MMX instruction set. In addition, AMD will bring the existing MMX capability on par with Intel's by implementing a second MMX execution unit, allowing superscalar execution of MMX instructions. In the second half of next year, AMD will introduce the second new K6 chip, the K6+ 3D. This chip adds to the K6 3D an on-chip 256K L2 cache.

As shown in Figure 1, the company will aggressively increase the K6 core frequency in 1998. Though the K6 is likely to lag behind Pentium II in core speed, AMD's plans show the K6 edging closer.

AMD-3D Combines MMX with Floating Point

In a bold move, matched in boldness by other non-Intel x86 vendors (see MPR 10/27/97, p. 22), AMD has strayed from strict Intel instruction-set conformance and will implement a new set of MMX-like instructions, called AMD-3D, for floating-point operands. These new operations are of AMD's own design and are not compatible with similar offerings from Cyrix and Centaur.

Like Intel's MMX instructions, the AMD-3D operations use the x86 floating-point registers for operand storage. Consequently, no OS changes are required to save and restore the register state for the new operations.

Favor did not disclose any details of these instructions, but we expect them to operate on pairs of single-precision (32-bit) FP operands, since the MMX registers are 64 bits wide, and thus compute two results per instruction. If both MMX units have AMD-3D capability, the new chips would produce four single-precision FP results per cycle. As with any MMX instruction, operands must be arranged and packed properly to achieve maximum performance.

For many 3D applications, including most games, single-precision FP computation is adequate. Thus, certain 3D geometry and lighting algorithms could exhibit a 2× to 4× speedup using the AMD-3D instructions.

The effective performance improvement at the user level—e.g., a 3D game's frame rate—will certainly be less, however, since even 3D geometry algorithms are not 100% single-precision FP code. Data movement, loop overhead, and other decision-making code will not be accelerated by AMD-3D. Still, for code in which the K6's out-of-order execution engine can overlap AMD-3D operations with integer code, the speedup could be 50% or more.

Though FP performance for code rewritten to use the AMD-3D instructions should reap a dramatic performance improvement, the performance of existing FP code will still lag that of Intel's chips (see MPR 9/15/97, p. 18), because the K6 lacks a pipelined FPU. The new K6 chips improve the existing FPU only slightly, adding the ability to overlap the execution of FXCH and FP computation instructions.

Given that Intel is expected to introduce FP MMX extensions, known as MMX2, in its Katmai chip in early 1999, AMD agrees that AMD-3D may have only a relatively short window of opportunity. The company plans to implement MMX2 soon after the details become available, and it plans to implement and support both MMX2 and AMD-3D in all its future x86 chips.

AMD is working on DirectX and OpenGL libraries and will make development tools available, including profiling and analysis applications to help software developers understand where to use AMD-3D to maximum advantage. Yet with an Intel standard looming less than a year away, it is hard to see where the value proposition lies for the ISVs, PC makers, and end users (see MPR 10/27/97, p. 35).

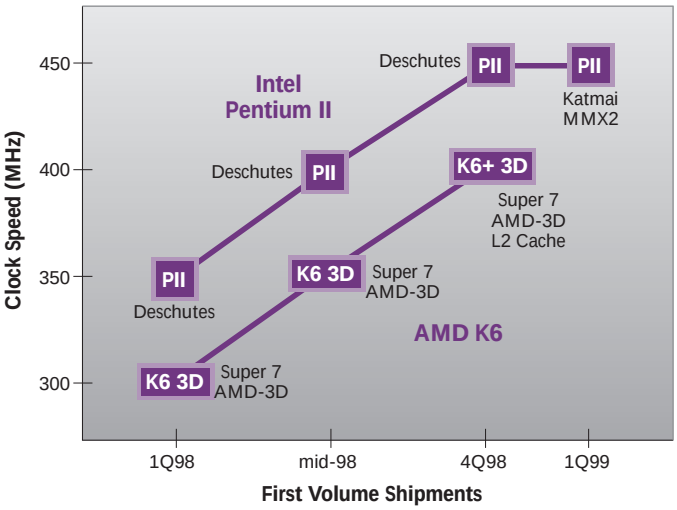

Figure 1. AMD's x86 roadmap shows steady improvement in clock speed to 400 MHz before the introduction of the next-generation K7 in the first half of 1999.

AMD, Cyrix Drive Socket 7

Without access to Intel's Slot 1 interface, AMD and the other x86 vendors are innovating to extend the life of the Pentium, or Socket 7, bus interface. The current Socket 7 interface is limited to 66-MHz operation, although some motherboards support 75-MHz and 83-MHz modes as well. Only Cyrix, however, has a chip specified for a 75-MHz bus.

Compared with Pentium II, a big weakness of competing x86 chips is inferior L2-cache performance. AMD will attack this problem first with the 100-MHz bus and later with the on-chip L2. For marketing reasons, AMD and the rest must have a 100-MHz bus to claim parity with Intel.

AMD is calling its 100-MHz specification Super 7. Cyrix has its own definition for 100-MHz operation, though AMD wants to work with Cyrix to arrive at a single standard. Since the companies haven't released details, it isn't clear how much the two specifications have in common.

Along with the 100-MHz external bus, AMD is planning to introduce AGP on K6 motherboards. To do this, AMD will enhance its own chip sets and is working with independent chip-set vendors to ensure compatibility.

To enable 100-MHz operation, the K6 3D and K6+ 3D will have slightly relaxed bus signal timing, and new chip sets must implement the tag store for the external cache on the north-bridge chip. This optimization saves the expense of a separate tag SRAM chip and ensures high-speed access to the tags. The new chip sets will support up to a 2M L2 cache; as with current Socket 7 boards, the L2 will be direct mapped.

Unlike Cyrix, AMD has no plans to support intermediate bus speeds between 66 and 100 MHz. The company expects 66- and 100-MHz motherboard designs to co-exist in the market for a while—until the price difference between 66- and 100-MHz cache SRAMs disappears—and the new K6 chips will work with either bus speed. To reach 400 MHz with a 66-MHz bus, the K6 chips will support clock multipliers up to 6×.

On-Chip L2 Operates at Processor Speed

AMD-3D will first appear in a CPU with just a conventional on-chip cache, but in 2H98, AMD will introduce the K6+ 3D with a 256K on-chip L2. Favor provided few details, but he did reveal that the cache operates at the full speed of the processor, interfaces to the CPU core through a dedicated 64-bit backside bus, and is four-way set-associative.

Using the K6+ 3D with integrated L2, the external cache, if present, becomes an L3. The majority of external caches on Socket 7 motherboards are limited to 512K, but with a 256K on-chip L2, such a small external cache will be largely ineffective and possibly even detrimental to performance. AMD will thus have chip sets supporting large external caches of at least 1M with plans for 2M. But for low-cost or space-constrained designs, especially notebooks, the large on-chip L2 of the K6+ 3D should deliver good performance without any external cache at all.

The advantage of an on-chip L2 at the high clock rates AMD will offer is significant. This cache will operate at the speed of the processor core, compared with one-half core speed, as implemented in Pentium II, or one-third or less in existing Socket 7 designs. Even with the 100-MHz Super 7 (see sidebar), the on-chip L2 will operate three times faster than an external cache. In addition, the on-chip L2 breaks the single-bus model of Socket 7 and gives the K6 a dual-bus architecture without breaking compatibility with Socket 7. AMD believes the on-chip L2 will boost K6+ 3D performance past that of Pentium II at the same core frequency.

Die Size Remains Modest

Counting the cache tag and data arrays plus control logic, the on-chip L2 of the K6+ 3D adds about 12 million transistors to the K6 3D core. Today's K6 uses 8.8 million transistors, with the K6 3D coming in at 9.3 million; the K6+ 3D will rack up an impressive 21.3 million.

The 0.25-micron K6 3D die will be 81 mm^2, half the size of the current 0.35-micron K6. Even though 0.25-micron cache cells pack extremely densely, the K6+ 3D die will grow by two-thirds to 135 mm^2. While this is significantly larger than the K6 3D, it is still easily manufacturable in high volume and is 17% smaller than the 0.35-micron K6.

The MDR Cost Model estimates the K6 3D will cost about $45 to build, only a few dollars more than the current K6 in the same 0.25-micron process. Thus, AMD will be able to price the K6 and the K6 3D essentially equally, allowing the company to obsolete the older part and aggressively push its AMD-3D instructions into the market.

We estimate the K6+ 3D, which should compete well with Pentium II at equivalent clock rates, will cost about $65 to build. Since the estimated cost of the Deschutes implementation of Pentium II is about $80 including the L2 cache, AMD should be able to become more profitable while continuing its practice of undercutting Intel's prices.

AMD to Enter Mobile Market With New Parts

Despite the enhancements of the new K6 chips, AMD expects their power consumption to remain within the envelope tolerated by laptop PCs. While the company expects the 0.25-micron K6 chips to be adopted in the mobile market, it is reluctant to say anything specific about its market plan.

Perhaps the most attractive K6 for laptops is the K6+ 3D, with its large on-chip L2 cache. In a mobile design, the on-chip cache eliminates the need for an external cache, saving board space, cost, and power. AMD claims power consumption is minimally increased by the on-chip L2 and that system power is decreased by generating fewer off-chip

accesses. The only problem is that the chip is not expected to appear until 2H98.

Another issue in the mobile market is packaging. The K6 uses C4 solder-bump technology, which is incompatible with TAB/TCP packages. Consequently, the K6 cannot be pin compatible with Intel's mobile chips. BGA packaging is a natural solution, but there is no standard pinout.

Super 7 Addresses Most Issues

Intel's introduction of the patent-protected Slot 1 was a bombshell for AMD and Cyrix. Contractually prevented from offering Slot 1 processors (see MPR 1/22/96, p. 5), AMD has been forced to innovate, and Super 7 fulfills AMD's promise to address the weaknesses of Socket 7.

The biggest weakness of Socket 7 for mainstream desktop and laptop designs is the lack of a dedicated L2 cache bus and the performance of the external bus in general. The on-chip L2 of the K6+ 3D addresses the first issue, and the 100-MHz bus addresses the second.

Another weakness of Socket 7 compared with Slot 1 is in multiprocessor support, and AMD's Super 7 enhancements do not address this issue. The target for AMD's Super 7, however, is the mainstream uniprocessor desktop and laptop market, so this technical deficiency should not prove a major weakness unless Intel decides to make multiprocessor support in mainstream PCs a key marketing thrust.

A serious weakness of Super 7 is simply that it is not Slot 1. Supporting two processor/motherboard interfaces is more difficult for the industry than supporting a single standard. In the short term, this will be only a small problem, since Socket 7 is still a part of most vendors' product lines. But as Intel deemphasizes its own Socket 7 processors, the problem could grow.

MICHAEL MUSTACCHI

AMD Greg's Favor explains how the K6's new MMX instructions improve 3D performance.

AMD Pacing Intel

The K6 roadmap spells out AMD's plans to compete with Intel's developments. In the first half of 1998, AMD will be in full production with 0.25-micron technology. The K6 and K6 3D will both be available at 300 MHz, and both chips will support the 100-MHz Super 7 external bus. In the same timeframe, Pentium II will be available at 350 MHz with a 100-MHz bus, but without MMX2. Thus, the K6 will be 10% or 20% behind in clock speed, but AMD can use the AMD-3D capability as a marketing advantage if there is enough software that uses it.

In the middle of next year, the K6 3D will be up to 350 MHz, compared with 400 MHz for the Pentium II. The K6 line will probably remain in about the same relative position: parity in bus speed and about 15% behind in raw clock speed, but with an advantage in 3D performance.

Price & Availability

AMD expects the K6 3D to ship in 1H98 and the K6+ 3D to ship in 2H98. No pricing was announced. For more information on the K6 family, contact AMD on the Web at *www.amd.com/K6*.

By the end of the year, Intel may implement its MMX2, potentially erasing the advantage of AMD-3D. At the same time, AMD plans to have 400-MHz K6+ 3D chips ready to compete with Intel's 450-MHz Pentium II. At this time, AMD is faced with another marketing challenge: how to communicate the benefits of the on-chip, full-speed 256K L2 cache compared with Pentium II's half-speed but larger 512K L2. Also by the end of the year, Intel will have Slot-2 Pentium II processors with full-speed 1M L2 caches, but these processors are likely to be expensive products intended for high-end servers and workstations, uncompetitive with the relatively low-cost K6 line.

Schedule Limits New Features

With the K6 3D and K6+ 3D, AMD appears to have been driven by schedule as much as by technical innovation. The schedule and die-size impacts of the K6 3D are minimal but gain the company some marketing advantage. AMD did not, however, fix the deficient scalar FP performance of the K6, citing schedule impact and claiming that AMD-3D addresses the most important uses of FP: 3D graphics.

Improved scalar FP will not come until the K7, a full redesign, appears in 1H99. Sanders said nothing about the K7 core but disclosed that the chip will be packaged in a module that is physically identical to a Slot 1 processor. Instead of using the P6 bus protocols, which would violate AMD's cross-license agreement with Intel, the K7 bus will be compatible with Digital's 21264 system bus (see MPR 10/28/96, p. 11). This 64-bit bus is designed to run at speeds of up to 333 MHz, offering more than three times the sustainable memory bandwidth of even a 100-MHz Slot 1. Sanders claimed this design will allow AMD to leapfrog Intel, but unless PC makers are willing to support higher clock speeds for the Digital bus than for Intel's P6 bus, the performance advantage will be minimal.

AMD's efforts do not threaten Intel's dominance of the PC processor market, but the new K6 chips will certainly compete very well in mainstream PC markets and may allow AMD to tap the mobile market as well. The new chips should fuel already high interest in the K6 throughout 1998 and, if AMD can overcome its current manufacturing problems, help the company gain new customers. M