Klamath Freezes, Direct RDRAM Cooks

Intel Shows Chilled 400-MHz Part at ISSCC '97 While DRAM Vendors Battle

by Jim Turley

Last month's International Solid-State Circuits Conference (ISSCC '97) offered its usual fare of clock circuits, PLLs, sigma-delta converters, and DRAM fabrication tricks. But amid the transistors and trenches, a number of interesting details emerged, including a faster-than-expected Klamath from Intel, a later-than-expected 21264 from Digital, and a larger-than-anyone-ever-expected 4-Gbit DRAM from NEC.

A panel discussion on DRAM technologies highlighted some fundamental differences of opinion among some of the major DRAM and logic makers about which interface standard will dominate the market at the close of this decade. Opinions were sharply divided on the prospects for synchronous DRAM, SyncLink, and the next-generation Rambus RDRAM, to be called Direct RDRAM. Despite Intel's choice of Direct RDRAM as the *de facto* standard for future PCs, other DRAM backers continue to whistle past the graveyard, devoted to their alternate approaches.

Fast Klamath Captures Media Attention

Probably the most talked-about presentation at the show was Intel's demonstration of the P6 chip known as Klamath *(see* 110201.PDF*)*, although the company never used that name, let alone the part's actual name, Pentium II.

The surprises started early as Intel presenter Mustafiz Choudhury changed his title slide from "A 300-MHz CMOS Processor with Multimedia Extensions" to "A 400-MHz…" He was careful to point out that the purpose of the "technology demonstration" was to show "the eventual potential of the P6 family" and that the chip was "not a product."

During the Q&A period, direct answers regarding the chip's performance and power consumption were avoided. When questioned about the chip's actual temperature and cooling method used to reach 400 MHz, Choudhury did admit that the chip was cooled with water to a temperature "colder than an ice cube," an apparent contradiction in solid-state physics.

Later that evening, Intel VPs privately demonstrated the chip in question. Purportedly running at 400 MHz, the system rested on a carefully draped table. Demonstrations of chilled commercial processors are not unknown, but neither are they commonplace. KryoTech showed its aggressively cooled Alpha and Pentium Pro systems at the most recent Comdex *(see* 1016MSB.PDF*)*. The company chills commercial-grade processors to $-40°C$ $(-40°F)$, boosting clock speeds for BiCMOS chips (like Pentium Pro) by 33% and pure CMOS parts by 50%. A 50% speed increase for a 266-MHz CMOS

Klamath would take the chip to 400 MHz—precisely the speed of Intel's technology demonstration.

Although it does not appear Intel was using KryoTech's cooling gear specifically, it is clear the chip was attached to a refrigeration device. A conventional Freon-charged unit may have been used to chill water, which was then piped into the demo system and over the Klamath module's surface.

In some sense, Intel's demo was no big deal. Semiconductor engineers have long known that CMOS circuits run faster at cooler temperatures and higher voltages. Intel merely demonstrated what most engineers already know.

By showing the system in a private suite after the show, Intel circumvented ISSCC rules regarding demonstrations. Even so, some observers grumbled, drawing an ethical distinction between heat sinks or fans that simply dissipate heat and refrigeration units that actively chill a part to far below room temperature.

Few Surprises in Processor Presentations

All told, this year's ISSCC included nine microprocessor presentations, with a mean clock frequency of 363 MHz. The presentation of Somerset's PowerPC G3 was in stark contrast to Intel's. Lumped in with a number of miscellaneous logic circuits and banished to the Saturday afternoon time slot, the presentation was poorly attended. Rather than show off an artificially accelerated processor, Motorola's Paul Reed described a 250-MHz, 0.25-micron implementation of the chip code-named Arthur *(see* 110203.PDF*)*. Given that the PowerPC 603e already runs at 240 MHz in a less-aggressive 0.33-micron process, Somerset's claims seem considerably more practical than Intel's.

At ISSCC, AMD discussed a 233-MHz K6, slightly faster than the company's earlier expectations for the chip in a 0.35-micron process. AMD described various esoterica of its K6 processor, including clock distribution, latch triggering, and testability features. The schedule for K6 appears to be on track, with production coming well before midyear. At 157 mm^2, the K6 has been trimmed by 5 mm^2 from AMD's estimates at the Microprocessor Forum *(see* 101406.PDF*)*. If AMD continues to execute this well for the next few months, the K6 should be a potent competitor to Pentium II by midyear.

Alas, Digital's execution has not been as good. The enormously complex 21264 still has not taped out, despite the company's predictions of 4Q96 tapeout *(see* 101402.PDF*)*. Tapeout is now planned for the end of March. Based on our rule of thumb of 12 months from tapeout to system shipments, volume production of the 21264 is now unlikely to happen before 2Q98. The die size of the 21264

has been fixed at 314 mm², meeting Digital's original expectations of being slightly larger than the 298-mm² 21164.

Also on the Alpha front, Mitsubishi's 21164PC was described as running at 550 MHz, right on target with Digital's earlier speed estimates. One interesting aspect of the 21164PC's design emerged: the cache is scanned dynamically at power up by BIST firmware, when bad rows are replaced with redundant rows. Most vendors catch hard defects in manufacturing during wafer probing, where failed cache rows are replaced via laser trimming. With its dynamic self-repair mechanism, Mitsubishi can avoid lengthy and expensive wafer testing and substantially reduce its testing costs.

Exponential's presentation on the x704 held few surprises, although one fabrication detail did slip out. The company's fab partner is using silicon-on-insulator (SOI) technology, a rarely used technique, to boost speed by increasing the dielectric isolation of the BiCMOS transistors. Two other SOI devices showed up at ISSCC, a DRAM from Mitsubishi and an ATM switch from NTT.

IBM Spins Single-Chip System/390 Processor

IBM described a single-chip version of the processor from its System/390 minicomputers that runs at 400 MHz. With 7.8 million transistors, the single-chip S/390 processor has only about half the transistor count of IBM's single-chip POWER processor, the P2SC *(see* 101104.PDF*)*, although at 300 mm² it's almost as big. The new chip is built using IBM's 0.27-micron CMOS-6S process *(see* 101203.PDF*)*, the same process used for the P2SC.

Internally, the S/390 chip maintains an identical pair of instruction units, fixed-point units, and floating-point units, which it runs in parallel. All instructions are executed in both units simultaneously and then checked before being committed to the ECC-protected register file. If there's a discrepancy in the results, the S/390 initiates error processing. Like its mainframe predecessors, this design allows the chip to recover from soft hardware errors.

Multimedia Chips Proliferate

The number of papers dedicated to media accelerators and processors attests to the growing importance of these new chips and the fundamental differences between them and conventional microprocessors. A total of nine media-related devices were described, including:

- An MPEG-2 video encoder from Philips
- NEC's MPEG-2 encoder
- A "video DSP" from Sony
- A French motion-estimation chip-set project
- An IDCT circuit for HDTV from LG Electronics
- An MPEG-2 decoder from LSI Logic
- Mitsubishi's D30V chip *(see* 101601.PDF*)*
- A 23-GOPS video processor from Matsushita
- A 72-mm² processor for compressing movies

NEC Chip Is Smart About Motion

NEC described a CMOS MPEG-2 encoder IC that shows more intelligence than most. The chip has 3.1 million transistors on a 155-mm² die and uses feedback from previous frames to assist with motion estimation. If, for example, objects in previous frames were moving from left to right, the chip adaptively shifts its search window in that direction. The search window can shift on each picture cycle based on the history of motion vectors. NEC's goal was to reduce power consumption and simplify the computing resources required for effective image compression.

Sony's video DSP presentation described a "2-RISC MIMD 6-PE SIMD" architecture for real-time MPEG-2 encoding and decoding. This odd design incorporates a RISC core for pixel processing and a second CPU core for flow control; in addition, a vector processor with six processing elements is assigned to the six blocks of a macro block in 4:2:0 format (hence, 2-RISC MIMD, 6-PE SIMD). As elaborate as the part is, it still requires an external motion-estimation chip, which the NEC part does not.

Sony has correctly determined that current conventional microprocessors, even those with multimedia instruction-set extensions, cannot perform MPEG-2 decoding (much less encoding) in real time. Despite some vendors' claims to the contrary, it now looks as though software-only 30-fps DVD playback (which relies on MPEG-2 for video encoding) will not be possible on a 233-MHz Pentium II processor. For the time being, PC-based MPEG-2 playback will still need at least some hardware assistance.

DRAM Devices Go Nonlinear

The award for most capacious DRAM goes to NEC for its huge 4-Gbit device. At a sprawling 986 mm², even in 0.15-micron CMOS, the device fairly groans under its own weight. A diagonal measurement of more than 44 mm (1.75 inches) makes one wonder what reticle NEC must be using to fabricate this behemoth, shown in Figure 1.
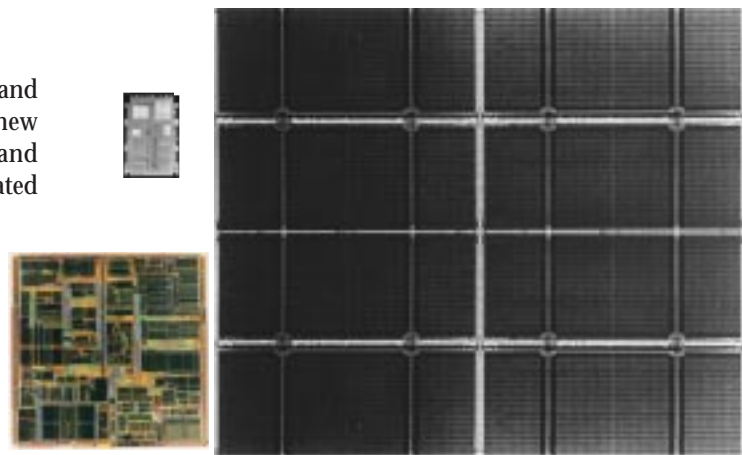

Figure 1. Die photos at a constant 2:1 scale compare NEC's (nonworking) 4-Gbit DRAM, at 986 mm², with Pentium II (203 mm²) and IBM's PowerPC 401GF (22 mm²).

More interesting than the device's fabrication characteristics is its design. To achieve 4-Gbit capacity, NEC used four-level bit encoding. Each "bit" cell stores four voltage levels, corresponding to two bits. Voltage levels at the rails (V_{CC} or ground) represent 11 and 00; 01 and 10 are represented by storing 1/3 or 2/3 the level of V_{CC}. Special sense amps distinguish among the four voltage levels and produce the proper binary results. To store two bits per cell, each cell must have three times the capacitance of a normal DRAM cell, according to NEC, which the company achieved with a special high-dielectric-constant material.

Unfortunately, the massive chip is massively bogus. NEC's 4-Gbit device is not close to working. The company can, in fact, get very few of its four-level cells to work reliably but chose to put two billion of them on a die anyway.

TI discussed a new DRAM design under development that places a sense amp at both ends of each bank of DRAM cells. The dual amps allow reads to occur from one sense amp while the other is precharging, so the DRAM is never off line due to precharge. In its current configuration, TI's development chip is built on a pure logic process, so each cell measures 33 μm^2, or about as large as a 6T SRAM cell. The dual-amp design could prove valuable in the future because it eliminates much of the dead time usually lost to RAS-precharge delays.

Rambus made a presentation on higher-bandwidth memories, but it did not involve Direct RDRAM. Using enhanced delay-locked loop (DLL) circuits and a fairly mundane 0.38-micron process, the company achieved data rates of 1 GHz per pin. With some further enhancements, Rambus expects to exceed this speed; a 32-bit version of such a device could move more than 5 Gbytes/sec. Clearly, much of the technology Rambus discussed at ISSCC could be used in future Direct RDRAM devices.

DRAM Turns Competitive

The DRAM sessions concluded with a panel at the end of the day. Topics strayed from the future of synchronous DRAM (SDRAM) to Rambus license fees to the official name (Direct RDRAM) of the next-generation Intel/Rambus interface *(see* 1017MSB.PDF*)*.

The panel agreed that, regardless of the underlying DRAM technology or interface method, normal FR-4 printed-circuit boards reach their limits at about 1.2 GHz. Beyond that, higher bandwidth must be achieved through wider, not faster, buses.

Some fundamental disagreements emerged over the issue of bandwidth vs. latency. Packet-based protocols, such as Rambus's, offer better theoretical bandwidth but longer initial latencies. High bandwidth is preferred for systems with predictable access patterns, while low latency helps systems with unpredictable accesses.

VLSI's Desi Rhoden seriously diminished his credibility with claims that SDRAMs have inherently better cores than any packet RAM (i.e., RDRAM) and that SyncLink will

definitely become the next major DRAM technology because it has more backers (16) than does Direct RDRAM, which has only two. Of course, those two happen to be Rambus and Intel.

On the other side of the debate, the Rambus partisans also made their share of blunders. Intel's Peter McWilliams argued that SDRAMs cannot be made to work at 133 MHz, which did nothing to sway members of the audience already running them at 166 MHz. Craig Hampel of Rambus unintentionally convulsed the audience in derisive laughter with the statement that "Direct RDRAM is an open standard," presumably in reference to the unpublished and proprietary Direct RDRAM specifications.

The issue of Rambus's license fees came up. The cynical point of view held that since Intel anointed Direct RDRAM as the interface of choice for the one system that consumes most of the planet's DRAM production, the sole owner of that technology, Rambus, has the DRAM industry over a barrel. The obvious question arose: What will prevent Rambus from gouging the world's DRAM vendors?

Apart from platitudes from the Intel and Rambus representatives regarding their altruistic motives, some reasoning emerged. To wit: it is not in Intel's best interest to stifle DRAM production; indeed, the company's strategic goals are best served by decreasing, not increasing, the price of personal computers. Intel would not be advocating Direct RDRAM if the company felt the devices would be significantly more expensive than synchronous DRAM or the other alternatives.

A Good Glimpse at Basic Technology

While the 400-MHz Pentium II demo captured most of the popular attention, it proved little about the part or Intel's roadmap for it. Even with obviously excessive thermal-management measures, the chip ran just barely (3%) faster than the average CPU speed at ISSCC. Without the ice, Pentium II would have been among the slowest processors at the show.

The proliferation of media processors from Japanese vendors underscores the growing importance of these chips in new consumer-electronics items and as adjuncts to mid-range PCs. The presentations also illustrated how the same problem can be attacked in many different ways, with varying levels of programmability and flexibility.

ISSCC has historically appealed to the down-and-dirty circuit designer, and the show still caters to researchers and academics pushing the boundaries of solid-state physics and fabrication. Most of the devices presented last month are still a gleam in some vendor's eye, but many are nearing production this year. The microprocessors, in particular, are nearly all destined for volume in 1997. As always, it's interesting to contrast the research projects with the almost-there products and marvel as one becomes the other. ◨

Printed proceedings of ISSCC '97 are available for $125 through the IEEE at www.sscs.org/isscc/digest.htm.