

## Digital Sells Its Chip Business

### *Intel Gets Fab, StrongArm in Settlement of Legal Battle*

by Linley Gwennap

As part of an agreement to settle the patent litigation between the two companies, Intel has agreed to buy Digital's entire semiconductor operation, including its Hudson (Mass.) fab, for approximately \$700 million. Intel will also pay Digital a smaller amount to set up a 10-year patent cross-license between the two companies, effectively ending the litigation. The deal reportedly improves Digital's discounts on future purchases of Intel processors as well.

Under the agreement, Intel will build Alpha chips for Digital, since the latter will no longer have a fab. Digital will continue to be responsible for developing and marketing Alpha microprocessors; Intel's role will be solely as a foundry. But as part of the deal, Digital has endorsed Intel's IA-64 technology, and we believe that, over time, IA-64 will supplant Alpha within Digital's product offerings.

In addition to the fab, Intel gets Digital's non-Alpha chip business, including the company's networking chips and its StrongArm microprocessor family; Digital will no longer market these products. Intel says it will offer jobs to all affected Digital employees, but it isn't clear how long it will maintain these product lines. Intel plans to fit out the Hudson fab for its new 0.25-micron process and to manufacture a variety of Intel products there, but it must await government approval, which could take months.

#### Alpha Products Remain on Track—For Now

Digital claims it remains committed to the Alpha architecture and that the Intel deal has no effect on its Alpha plans. Like Silicon Graphics and Sun, Digital will simply have its processors built by an external fab.

Relying on Intel as a foundry could have advantages. In the short term, the Hudson fab will continue to run Digital's 0.35-micron process, and in the future, Digital's 0.25-micron process. Under Intel's management, the fab will also run Intel's 0.25-micron process, which is slightly different from

Digital's (see MPR 9/16/97, p. 11). Intel believes that both 0.25-micron processes will run on the same production line, however, so it will simply switch from one process to the other, depending on the type of chip being manufactured.

In the 0.18-micron generation, Digital will design its chips for Intel's process, eliminating this bifurcation. Digital has often lagged Intel by as much as a year in deploying a new process generation; the deal could allow Digital to ship 0.18-micron Alpha processors well before it could have without Intel. On the other hand, Intel's process may lack some of the performance-enhancing features included in Digital's current designs, but Intel's 0.18-micron process will unquestionably be faster than Digital's 0.25-micron process.

The downside for Digital may be a lack of responsiveness from the fab. Today, the Hudson fab can tweak the process to achieve maximum clock speed and select a handful of parts from the far end of the yield curve to fill Digital's fastest speed grades. Under Intel, the fab will be managed to achieve high volume, not the maximum possible performance, and exceptions for the Alpha parts may be more difficult to make. Digital says it has contractual guarantees that Intel will produce Alpha chips at the same speeds they are today, but we would not be surprised if the transition made it more difficult for Digital to achieve industry-leading clock speeds in the future.

#### Interest in Alpha Will Wane

In the long term, we suspect Digital will phase out its Alpha line. Once the company begins selling Merced systems in 1999, interest in the Alpha boxes is likely to wane. Digital said it will port Digital Unix to IA-64, and Microsoft has already announced plans for Windows NT on Merced, so nearly all of Digital's workstation customers and at least half of its server customers will be able to move easily to the new Intel-based systems. (Digital's other customers use OpenVMS, which will probably remain on Alpha only.)

*Continued on page 6*

## AT A GLANCE

<b>Digital Sells Its Chip Business</b> . . . . .	<b>1</b>
Intel will purchase Digital's semiconductor business, including its only fab, for \$700 million to help settle the companies' lawsuits. Although Intel will build Alpha chips for Digital, the computer maker is likely to move from Alpha to IA-64 over time.	
<b>Editorial: Software Spiral, Intel Profits Both Stall</b> . . . . .	<b>3</b>
A lack of compelling new software for high-end PCs has caused a shift in the market toward low-end boxes, reducing Intel's profits.	
<b>Most Significant Bits</b> . . . . .	<b>4</b>
Intel plans low-cost Pentium II products; Cyrix readies MediaGX with MMX; S3 brings 3D to business desktops; Bug makes some Pentium systems vulnerable; IBM speeds P2SC, deploys first 64-bit chip; TI adds floating point to 'C62xx.	
<b>Embedded News</b> . . . . .	<b>8</b>
Motorola updates ColdFire core to V3; 68K-to-ColdFire software translator emerges; IDT R4700 hits 200 MHz, New core-logic support for M32R/D.	
<b>MicroJava Pushes Bytecode Performance</b> . . . . .	<b>9</b>
According to Sun, the MicroJava 701 will, when it ships in 2H98, execute Java bytecodes faster than any other processor while reducing memory demands, due to its native Java execution engine.	
<b>Siemens TriCore Revives CISC Techniques</b> . . . . .	<b>13</b>
TriCore's retro-CISC instruction set combines microcontroller, DSP, and CPU features in a single processor core.	
<b>Centaur Improves C6 With No Extra Cost</b> . . . . .	<b>17</b>
Even as it begins shipping its initial C6, IDT subsidiary Centaur has revealed plans to improve that part's performance, primarily on FP and MMX code, by 2Q98. New clock speeds will also boost the chip.	
<b>PA-8500's 1.5M Cache Aids Performance</b> . . . . .	<b>20</b>
HP's next processor will include a record 130 million transistors and will reduce system cost even as it improves performance.	
<b>Gshare, "Agrees" Aid Branch Prediction</b> . . . . .	<b>22</b>
Two new branch-prediction algorithms will become widely used in microprocessors due to their minimal cost and improved accuracy.	
<b>IBM's Power3 to Replace P2SC</b> . . . . .	<b>23</b>
IBM has modified the failed PowerPC 620 processor to produce Power3, the first 64-bit PowerPC chip with high floating-point performance. Power3 will replace the P2SC in 2H98.	
<b>The Slater Perspective: RISC on the Desktop—Game Over</b> . . . . .	<b>28</b>
With Digital now joining HP, Compaq, IBM, SGI, and others in the IA-64 camp, RISC workstations are likely to disappear over time.	
<b>Recent IC Announcements</b> . . . . .	<b>29</b>
<b>Patent Watch</b> . . . . .	<b>30</b>
<b>Literature Watch</b> . . . . .	<b>31</b>
<b>Resources</b> . . . . .	<b>32</b>

MICROPROCESSOR  REPORT

## Founder and Editorial Director

Michael Slater  
mslater@mdr.zd.com

## Publisher and Editor in Chief

Linley Gwennap  
linley@mdr.zd.com

## Senior Editors

Jim Turley  
jturley@mdr.zd.comPeter Song  
psong@mdr.zd.com

## Senior PC Analyst

Peter N. Glaskowsky  
png@mdr.zd.com

Editorial Assistant: Arlene Lum

## Editorial Board

Dennis Allison	Rich Belgard
Brian Case	Jeff Deusch
Dave Epstein	Don Gaubatz
John Novitsky	Bernard Peuto
Nick Tredennick	John F. Wakerly

## Editorial Office

298 S. Sunnyvale Avenue  
Sunnyvale, CA 94086-6245

Phone: 408.328.3900 Fax: 408.737.2242

Microprocessor Report (ISSN 0899-9341) is published every three weeks, 17 issues per year. Rates are: N. America: \$595 per year, \$1,095 for two years. Europe: £450 per year, £795 for two years. Elsewhere: \$695 per year, \$1,295 for two years. Back issues are available.

## Published by

MICRODESIGN  
RESOURCESPresident: Peter Christy  
pchristy@mdr.zd.com

## Business Office

874 Gravenstein Hwy. So., Suite 14  
Sebastopol, CA 95472

Phone: 707.824.4004 Fax: 707.823.0504

Subscriptions: 707.824.4001

cs@mdr.zd.com

World Wide Web: [www.MDRonline.com](http://www.MDRonline.com)

Copyright ©1997, MicroDesign Resources. All rights reserved. No part of this newsletter may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior written permission.

Computer Press Award, Best Newsletter,  
Winner, 1993, 1994; Runner-Up, 1996

 Printed on recycled paper with soy ink.

# Software Spiral, Intel Profits Both Stall

*Without Compelling New Software, Demand for High-End PCs Weakens*



Intel recently spooked investors by reporting weak results for the third quarter and saying the fourth quarter wouldn't be much better, as we anticipated (see MPR 7/14/97, p. 1). The principal cause is a precipitous drop in the average selling price (ASP) of Intel processors, due to a shift in PC sales toward less expensive systems. Many have characterized this shift as an increase in demand for sub-\$1,000 PCs, but I see the other side of the coin: weak demand for more expensive systems.

Intel's business model is built around continually offering more performance at the same price points. Its frequent improvements in technology drive this supply side of the equation, but Intel counts on corresponding increases in the performance requirements of popular software to build demand for its faster processors. Without this increase in demand, PC buyers might simply choose to buy the same processors quarter after quarter as their price declines.

That appears to be exactly what is happening today. Intel introduced the Pentium/MMX-200 in January of this year at a price of \$539. At the time, this chip was the fastest Intel processor for Windows 95 PCs (Pentium Pro being aimed at Windows NT systems). Pentium/MMX runs Windows 95 and most PC applications quite capably, including those that use the new MMX instructions.

Over the course of the year, the price of that chip has dropped to \$213, and it still handles current applications with few problems. The 166-MHz Pentium/MMX, which is only about 10% slower, sells for just \$112 and is starting to appear in sub-\$1,000 PCs. These powerful yet low-cost processors are spurring demand for inexpensive PCs.

Low-end PCs have also benefited from recent drops in DRAM and hard-drive prices. Despite the collapse of DRAM prices in 2Q96, memory prices have not stabilized and are now about an eighth of their 1Q96 level. Thus, even an inexpensive PC now comes with 32M of memory, more than enough for Windows 95 and typical applications. Similarly, it's hard to find a system with less than a gigabyte of disk, which is plenty even for obese Microsoft software.

Thus, there are few compelling reasons for PC buyers to acquire \$2,000 Pentium II systems with monster hard drives and boatloads of DRAM. These high-end systems appeal to the usual professionals that run Photoshop or AutoCAD all day, and to the dedicated 3D gamers that need (and can afford) the fastest machine on the block. But with the least-expensive Pentium II processor priced at \$401,

these systems are still too expensive for the mainstream consumer or business user.

What Intel needs is compelling software that requires the performance of Pentium II. God knows they've tried to promote CPU-hungry applications such as digital photography, videoconferencing, DVD playback, and voice recognition, but these applications haven't caught on, in many cases due to a lack of infrastructure. Digital cameras with reasonable picture quality are still too expensive, and current photo-editing software is too hard to use. Few PCs have enough bandwidth for reasonable videoconferencing. Most people want to watch movies on their TVs, not their PCs. And voice recognition hasn't been integrated into interesting applications.

The emergence of Microsoft's Windows 95 was a key factor in obsoleting the 486. The perpetually imminent release of Windows 98 won't help Intel, however, because it is just an incremental release that doesn't change the hardware requirements; like the current version of Windows, it should run fine on Pentium systems. In the corporate market, however, an ongoing shift toward the more powerful Windows NT may aid Intel's Pentium II push.

Another way to build demand for more expensive processors would be to introduce new features, such as the much-rumored MMX2. Sure, MMX2 will help only the handful of programs that use 3D graphics, but who cares? MMX was a big success even though the number of programs that actually use it is small. Intel's competitors all plan to offer MMX2-like features, some as early as 1Q98, but it looks as if Intel won't have MMX2 until 1999.

Technically, it's fairly simple to add the parallel floating-point operations said to be the key component of MMX2. The problem for Intel is finding a release vehicle. Klamath was the first P6 processor with MMX and was too early for MMX2. To avoid delays, Deschutes is as similar as possible to Klamath. Katmai, the next train, doesn't leave the station until early 1999.

Thus, not enough customers are buying Pentium II systems, and Intel's ASP has sagged. In 1998, Intel plans to drop Pentium II prices so low that customers will have to buy them, but that won't do much to raise Intel's ASP. The company must hope that the infrastructure issues holding back emerging high-end applications are solved, or that some compelling new software emerges—soon. □



### ■ Intel Plans Low-Cost Pentium II Products

Intel has confirmed plans to offer two new versions of its Pentium II processor in 1998, both aimed at low-cost PCs. The first is a version of the module without L2 cache; this part is set to appear in mid-1998. The second product incorporates the L2 cache onto the processor chip. Because this part requires changing the CPU itself, it probably won't appear until late 1998.

With much of the growth in the PC market coming at the low end (see page 3), Intel wants to make sure Pentium II can participate at these price points. Even with the forthcoming 0.25-micron Deschutes, we estimate the cost of the module to be \$70, much more than the \$45 Pentium/MMX chip. Thus, selling the Deschutes module for the same \$100 price as a low-end Pentium/MMX would reduce Intel's margins—something the company doesn't want to do.

Eliminating the 512K cache and possibly trimming other packaging costs could shave \$5–\$10 from Pentium II's manufacturing cost, a significant margin improvement. These changes could be made while still allowing the cacheless module to plug into the same Slot 1 interface. Of course, a Pentium II with no L2 cache would have poorer performance, but probably no worse than a Pentium/MMX of the same clock speed on most applications. This performance would be adequate for low-end systems.

Integrating the L2 cache would solve the performance problem. Because the on-chip cache would presumably operate at the full CPU speed and could be set associative, it could be smaller than the current external cache—perhaps 128K or 256K—while providing similar performance. In a 0.25-micron process, a 256K cache would add about 30 mm<sup>2</sup> to the die, but the cost of the package could be reduced by eliminating the external cache interface.

Overall, adding the L2 cache on chip could increase the manufacturing cost of the processor by about \$5–\$10. But by getting rid of the external cache, this module would cost less than the current module with similar performance. The on-chip cache would also be ideal for notebooks, since it saves space and reduces power consumption. Thus, this design could become widely used in notebooks and mainstream desktops in 1H99.

The objectives of this strategy are clear: eliminate Socket 7 from Intel's product line as quickly as possible without giving up on the low-cost PC market. Because Pentium II uses Intel's proprietary Slot 1 technology, other x86 vendors plan to stick with the more open Socket 7 throughout 1998. By moving completely to Slot 1 by mid-1998, Intel can pull the rug out from under its competitors, forcing PC makers to maintain a different motherboard solely for the non-Intel parts. While Socket 7 will certainly remain viable after Intel ceases to use it, this move will make it more difficult for Intel's competitors to gain PC design wins. —L.G.

### ■ Cyrix Readies MediaGX With MMX

Cyrix is preparing an enhanced version of its low-end MediaGX processor (see MPR 3/10/97, p. 1) that adds MMX capability. The new chip, code-named GXm (GX multimedia), retains the 5x86 integer core but adds the fast FPU and MMX unit from the 6x86MX. The new chip also upgrades the EDO DRAM controller to support 100-MHz SDRAM.

Cyrix expects the initial clock frequencies to be 200 and 233 MHz. The GXm will be built in the same 0.35-micron four-layer-metal technology as the current MediaGX chips, which are now shipping at speeds of up to 200 MHz. The jump to 233 MHz comes from circuit rework.

The GXm enables the MediaGX family to keep up with the advances in Intel's low-end offerings; it won't move the family upscale, but it should keep it from falling further behind. Without the addition of MMX, the MediaGX would have had a hard time holding any market position in 1998. The big boost will come with the MXi, due in 2H98, which is based on the Cayenne CPU core (see MPR 10/27/97, p. 22) and includes an advanced 3D graphics accelerator. —M.S.

### ■ S3 Brings 3D to Business Desktops

Combining its successful Virge 3D engine with a new 128-bit GUI accelerator, S3's new Trio3D provides the features business users need—excellent 2D performance and low cost—along with checkbox 3D acceleration and basic support for digital video. Trio3D is sampling now. When S3 ships this new chip in the first half of 1998, it will be priced at \$22 in 10,000-unit quantities.

With this announcement, S3 ([www.s3.com](http://www.s3.com)) debuts a new 336-pin BGA package it plans to use for future consumer and business 3D products. The larger package supports 125-MHz SGRAM and provides extra pins for expansion; we expect to see full AGP support and possibly a 128-bit local-memory interface on 1998 products. S3 has recently acquired a license to use Rambus memory technology, and we also expect to see Direct RDRAM support in future S3 chips. For use in existing Trio64 and Virge products, the company will also offer Trio3D in a pin-compatible 208-pin PQFP. —P.N.G.

### ■ Bug Makes Some Pentium Systems Vulnerable

Intel has acknowledged a bug in all Pentium/MMX and Pentium processors that could enable malicious users to crash unprotected Internet servers. If a "lock" prefix is applied to a CMPXCH8B mem64 instruction that invalidly uses a 32-bit register (instead of a 64-bit memory value) as a destination operand, the processor comes to a complete halt. The correct response would be to signal an invalid operand exception. Since this instruction encoding is invalid, there is no reason for any compiler to produce the instruction or for any application to use it.



Thus, PC users running standard applications should never encounter this problem. The significance of the bug comes from the fact that a user with a shell account at an Internet service provider with the privilege to execute machine code directly could run a script that triggers the bug. There would be no reason for doing so other than malice, but it makes any system that gives users the ability to run user code subject to being deliberately crashed. (Note that clever and malicious users can find many other ways to create havoc in such systems.)

Intel believes there are potential workarounds, and it hopes to release further details by mid-November. These workarounds would probably be in the operating system and would only need to be applied on servers and other systems providing remote access; most PC users would not need to be concerned. Intel has no plans to recall processors or offer replacement parts due to this bug. —*M.S.*

### ■ IBM Speeds P2SC, Deploys First 64-Bit Chip

Two new systems from IBM improve the performance of the RS/6000 family ([www.rs6000.ibm.com](http://www.rs6000.ibm.com)) in two different ways. The Model 397 uses a 160-MHz version of the P2SC processor (see MPR 8/26/96, p. 14) that offers 20% better performance than the previous P2SC. The company also rolled out its first 64-bit RS/6000 system, the Model S70.

The Model 397 is rated at 7.1 SPECint95 (base) and a stellar 22.4 SPECfp95 (base). Although the integer score is no better than Pentium/MMX's, the floating-point score trails only HP's 236-MHz PA-8200 and Digital's 21164-600 among currently shipping systems.

The new processor is identical to the original P2SC except for a shrink to IBM's CMOS-6S2 process (see MPR 9/16/96, p. 11), a 0.25-micron hybrid process with metal layers similar to those of a 0.35-micron process. The new process improves the clock speed by about 20% while reducing the die size from 355 mm<sup>2</sup> to 256 mm<sup>2</sup>. Due to its enormous die and 1,088-pin package, the older P2SC had the highest manufacturing cost of any current microprocessor, according to the MDR Cost Model; the new version reduces the estimated cost from \$375 to \$290, better but still more than any shipping CPU.

The Model 397 includes the 160-MHz P2SC, 128M of memory, and a 4.5G hard drive (but no monitor or graphics acceleration) for a list price of \$29,900. IBM will also offer the faster PS2C in its SP line of parallel supercomputers.

For nearly two years, IBM has been shipping 64-bit AS/400 systems based on two "PowerPC AS" processors known as the A10 and A30 (see MPR 7/31/95, p. 15); these chips implemented the PowerPC instruction set along with special extensions to support older AS/400 software. The PowerPC 620 was supposed to be the first true PowerPC chip with 64-bit capabilities, but that processor floundered, and IBM now says it will not ship any 620-based products.

Instead, the company is selling a system based on a new processor called the RS64. This chip was derived from the

A10 but eliminates the AS/400-specific instructions. The chip acts as a stopgap solution until the 64-bit Power3 (see page 23) arrives next year.

The new S70 system uses a 125-MHz RS64 processor with 64K of instruction cache and 64K of data cache on chip. The standard configuration includes four CPUs, each with 4M of external cache, plus 512M of memory and 4.5G of hard disk, all at a list price of \$125,000. It comes with AIX 4.3, a new version of IBM's Unix operating system that includes full 64-bit support.


IBM's RS/6000 is the last of the major RISC product lines to gain a 64-bit processor. Digital and Silicon Graphics have been shipping 64-bit systems for years and offer this capability in all of their current systems; Sun and HP came later to the party but now have 64-bit processors in all but their low-end systems. Although the number of applications that take advantage of 64-bit support remains small, 64-bit addressing provides a significant performance boost when working with large data bases and large scientific data sets. Adding this capability will allow IBM to better compete in these high-end applications. —*L.G.*

### ■ TI Adds Floating Point to 'C62xx

Revving up its floating-point engine, Texas Instruments disclosed plans for the 320C67xx DSP, which adds floating-point capability to its EPIC-like 320C62xx DSP (see MPR 2/17/97, p. 14). The 'C67xx uses the same basic instruction set as the fixed-point 'C62xx, which will allow programmers to more easily move their code from one family to the other. Surprisingly, TI is the first DSP vendor to offer a common instruction set for both fixed- and floating-point DSP chips.

The 'C67xx core will be capable of six 32-bit floating-point operations per cycle, resulting in a peak execution rate of 1 GFLOPS at 167 MHz. This clock speed is slightly slower than that of the 'C62xx when built in the same 0.25-micron (drawn) five-layer-metal process. Still, this level of floating-point performance is well ahead of what is available in today's DSP chips, as TI has let its floating-point DSPs stagnate over the past few years.

The device is still a ways away: production parts won't ship until early 1999. TI says it will provide a follow-on part by the end of the decade that delivers 3 GFLOPS. The company did not announce specific products or pricing but said the 'C67xx parts will be priced about the same as current 'C3x and 'C4x chips, including some priced below \$50.

Many DSP designers prototype algorithms in floating-point arithmetic but use fixed-point chips for volume production, due to their lower cost. The 'C67xx will make it easier for designers to make this move. In addition, its high performance will improve the capabilities of advanced signal-processing applications such as voice recognition, cellular-telephone base stations, radar, and finite-element analysis. The combination of the 'C62xx and 'C67xx should give Texas Instruments the undisputed performance lead for both fixed- and floating-point DSP applications. —*L.G.* 

## Digital

*Continued from page 1*

Digital insists its Alpha chips will maintain a performance edge over Merced on at least some applications, and this may be true at the very high end. But we expect Merced to match or beat Alpha's performance on most applications (see MPR 10/27/97, p. 1), leaving little space for Alpha. With Merced in the mix, sales of Alpha systems have nowhere to go but down from their already modest levels.

Digital could ease its customers' conversions by developing an Alpha-to-IA-64 binary translator. This would be much simpler than the company's current x86-to-Alpha product, FX!32 (see MPR 3/5/96, p. 11), because of the more straightforward nature of Alpha code. Digital executives would not commit to developing such a product, but the company clearly has the expertise to do so.

Digital denies any plans to phase out Alpha and in fact has not given up its quest to move Alpha into the high-volume PC market. But the company's inability to deliver the low-cost 21164PC and the market's unenthusiastic response to FX!32 have left Digital unable to sign any significant PC makers for Alpha, despite the support of chip makers Mitsubishi and Samsung. Without Digital Semiconductor, the company will have no sales force dedicated to Alpha chips, making it even more difficult to find new chip customers.

Intel must be counting on Digital to move away from Alpha: the company has agreed to provide leading-edge fabrication technology to a chip that will be Merced's toughest competitor on performance, an arrangement that doesn't make sense in the long term. Although Digital CEO Robert Palmer has been an Alpha champion, Digital's board of directors is reportedly in favor of jettisoning Alpha, a stance that may have enabled the settlement (see sidebar, next page).

### StrongArm May Slip Through Intel's Grasp

Since Digital dumped its entire semiconductor operation in Intel's lap, Intel is now trying to decide what to do with the various pieces. Digital's lesser-known products include PCI-to-PCI bridge chips and 10/100-Mbit Ethernet interfaces, some of which have become popular in some circles. These devices should easily fit into Intel's existing product lines and support Intel's objective of supplying silicon for high-end systems based on Intel processors.

StrongArm is a different story. This inexpensive chip (see MPR 2/12/96, p. 1) is one of the fastest embedded processors available, but its power consumption is low enough for many portable devices. As such, it is a perfect complement for Intel's current embedded offering, the i960, which is expensive, slow, and power hungry. Whereas the i960 made its name in laser printers and other office equipment, StrongArm is ideal for PDAs, set-top boxes, and network computers.

There's the rub. Intel's corporate strategy is that every product should enhance the PC. Even the i960 has recently

been recast as an I/O processor for high-end x86 servers. Adding another embedded processor that isn't PC compatible could be a needless distraction for Intel.

Worse yet, StrongArm conflicts with some of Intel's existing PC-based initiatives. For example, Intel is promoting x86-based reference designs for network computers and set-top boxes, products where StrongArm-based chips such as the SA-1100 (see MPR 9/15/97, p. 1) are technically superior to the x86 offerings.

A third strike against StrongArm is its use of a non-Intel architecture. Among the assets being acquired by Intel is Digital's ARM license; although Intel says some details must be worked out, we expect ARM would be happy to let Intel build StrongArm chips. But Intel has never used an instruction set it didn't own, and NIH (not invented here) is the prevailing mentality at the microprocessor giant.

If Intel decides it is interested in making an aggressive move into emerging consumer-electronics markets, however, StrongArm is the perfect vehicle. An Intel StrongArm (two words that seem natural together) could dominate the market for Windows CE devices, creating a new Wintel axis. These products could offer incremental revenue growth to the company, albeit with smaller margins than for x86 processors. Intel must decide whether to maintain its shaky position that the PC is the solution to all problems, or take advantage of a fortuitous opportunity to launch a new processor product line.

### New Fab Could Aid Intel's Graphics Initiative

Intel plans its fab capacity years in advance, so a sudden decision to buy a fab is unusual. In fact, Intel has never purchased a fab before, although it has relied on external foundries for cache and other peripheral chips. The initial idea for Intel to purchase the Hudson fab probably came from Digital, but Intel could have simply resold the excess fab capacity or used it as-is for older products (chip sets, etc.). Instead, Intel says it hopes to ramp the Hudson fab to full capacity after upgrading it to the company's leading-edge 0.25-micron process.

One theory is that Intel underestimated the demand for its 0.25-micron capacity during its initial planning cycle and needed a quick fix. Certainly, the company has been capacity constrained recently (see MPR 4/21/97, p. 3), but those limitations appear to be easing already. The Hudson fab probably won't start producing 0.25-micron Intel chips before 3Q98, too late to help accelerate the conversion to Pentium II.

Another use for the new fab could be for Intel's graphics initiative. Intel's initial 0.25-micron fab plans may have neglected graphics chips, assuming they could be built on older fabs. It is now apparent that competitive 3D graphics chips will require 0.25-micron technology. If Intel is able to grab 10% of the PC graphics market in 1999, these chips could consume up to half the capacity of the Hudson plant, with at least some of the rest devoted to Digital's needs.

In any case, the new plant gives Intel an enormous amount of 0.25-micron capacity. The plan now includes five

fabs running the 0.25-micron process (Santa Clara, Chandler, Albuquerque, Leixlip, and Hudson), compared with just three for the current 0.35-micron process. One benefit of the Hudson purchase is that Intel can delay the build-out of its planned Ft. Worth (Texas) fab from 1999 to 2000. That plant was originally planned as a 0.25-micron fab but will now start at the 0.18-micron level.


### Getting Approval May Be Dickey

The settlement is on hold until it is approved by both the U.S. Federal Trade Commission (FTC) and the judges presiding over the suits. To gain FTC approval, the parties must show that the new arrangement does not diminish competition in the microprocessor market. If Digital were to admit that it plans to phase out Alpha, the FTC would almost certainly refuse to let the deal go through.

Thus, the dichotomy between Digital's public statements and private comments. In particular, Palmer and other Digital executives have taken a strong public position that Alpha will continue well into the future and is actually helped by this deal. Of course, this stance is necessary to protect Digital's Alpha business until Merced systems are available, but it also paints a picture that the FTC wants to see. Similarly, Intel will undoubtedly refuse to make any negative comments about the future of StrongArm until the deal is complete, a process that the companies expect will take up to six months.

The FTC is already investigating Intel's business practices, in part because of the heavy-handed way it originally responded to the Digital suit, and has a separate investigation of Intel's agreement to purchase Chips and Technologies (see MPR 8/25/97, p. 4). Its examination of the new agreement may take several months, but given the agency's unwillingness to hold up such deals in the past, this one seems likely to ultimately get a green light.

Although Intel will not admit guilt, of course, the form of the settlement implies the company was concerned that Digital could prove patent infringement. Typically, patent cross-license deals are not royalty bearing; both companies simply exchange patents. In this case, however, Intel reportedly agreed to pay Digital \$200 million over four years in addition providing access to its patents. Intel has also granted Digital the prestigious Tier 1 discount status; because of its moderate x86 volumes, Digital has been in Tier 2, although the companies would not confirm any such details of the agreement. Intel doesn't want any royalty payment to be reported, since it might encourage other lawsuits.

The settlement clearly weakens a threat to Intel's product line and strengthens IA-64's hold on the high-end system market. Digital characterizes the change as a win for Alpha, but in fact it is likely to ultimately remove Digital from the processor business, a tragic fall for the company that invented the minicomputer more than 30 years ago. And so, we move one step closer to a world in which all significant computers are built using Intel microprocessors. 

## The Making of The Deal

When Digital first sued Intel (see MPR 6/2/97, p. 26), Intel's motivations were obvious: make the suit go away before it jeopardized the enormous revenue stream from Intel's Pentium and Pentium II products. Yet the company was eager to avoid appearing to make a large payment to settle the suit, since that might seem to be giving in to patent blackmail and thus encourage similar suits.

Digital's motivations for its suit were less apparent. Obtaining some payment from Intel is clearly a benefit for the financially weak company. The patent cross-license gives Digital more flexibility in its future processor designs and settles Intel's counterclaim of patent infringement.

At some point in the talks, Digital's semiconductor operations came into play. Digital has long maintained its own fabs to build high-performance chips for its computer systems, but the Hudson fab was too large for the modest number of chips that Digital currently requires. Digital set up its semiconductor business in 1993 to develop and sell chips on the open market, hoping to create enough demand to fill the fab. Unfortunately, neither the Alpha processors nor Digital's other chips became big sellers, leaving the semiconductor operation losing as much as \$100 million a year, according to one report.

Furthermore, Digital's board of directors, along with some executives in the systems group, had allegedly become disenchanted with Alpha. Despite its technical superiority, Alpha has not caused Digital's system sales to surge, and no other large computer vendor has adopted the architecture. Substituting Intel's IA-64 technology for Alpha would cut Digital's costs while providing access to industry-standard hardware and software.

Digital can't convert to IA-64 immediately, since Merced won't ship for two years, and Digital has a large installed base of Alpha customers. Thus, Intel had to commit to building Alpha chips during a potentially lengthy transition period. Sources indicate Intel is required to supply Alpha chips for up to seven years.

Thus, the agreement lets Digital get rid of its unprofitable semiconductor operation and focus on its primary mission of providing high-performance computer systems and service. The lawsuit had chilled Digital's relations with Intel, but now Digital has unfettered access to all of Intel's products again.

Intel has managed to both eliminate a threat to its revenue stream and bring Digital on board as an IA-64 customer, a move that, before the first IA-64 chip even ships, almost guarantees Merced will become the leading processor for high-end systems (see page 28). Intel also gains a useful fab for merely book value. Compared with Intel's \$8 billion cash hoard, the cost of settling the suit is, in the words of Intel president Craig Barrett, "not material."



### ■ Motorola Updates ColdFire Core to V3

As Motorola's midrange 32-bit approaches adolescence, it is undergoing some internal changes. Motorola has released details of ColdFire V3, a microarchitectural alteration that will allow future ColdFire chips to keep up with industry-wide advances in clock speed and performance. The first ColdFire 3 chips are expected to debut in 1H98.

To allow for faster clock rates, V3 extends ColdFire's pipeline by two stages. Instruction decoding and operand reads now take two cycles apiece. The instruction buffer has also been enlarged to hold eight instructions rather than 12 bytes (3–6 instructions). The longer pipeline means branch instructions actually take longer than before (in terms of clock cycles), but this penalty will be offset by V3's faster peak clock rate, which Motorola expects to reach 100 MHz.

To alleviate some of the branch penalty, ColdFire V3 implements a new form of branch "hinting." Programmers can use a new global control bit to reverse ColdFire's usual prediction for forward branches (i.e., taken vs. not taken). For software that knows which way branches are likely to go, this new feature can eliminate many mispredictions.

All the changes to the pipeline allow Motorola to support higher frequencies and clock multiplication for the first time with ColdFire. Whereas today's parts are limited to about 33 MHz, next-generation ColdFire parts will reach 90–100 MHz in the next 18 months, according to the company. Fabrication will also move to 0.35-micron technology from the quaint 0.8-micron methods used now.

Motorola expects that most ColdFire V3 parts will include a hardware multiply-accumulate (MAC) unit, which few ColdFire chips have now, and that they will support an integer divide instruction for the first time.

The company predicts that 90-MHz ColdFire V3 parts will be nearly 3× faster than current 33-MHz V2 parts—which is hardly surprising, given the nearly 3× difference in clock speed. The longer pipeline, clock doubling, and branch hinting merely allow ColdFire to keep up with its clock rate. Clock-for-clock, ColdFire performance will be unchanged.

Overall, Motorola's alterations to ColdFire are nothing spectacular; they merely keep ColdFire on the growth path the company outlined last year (see MPR 9/16/96, p. 1). The company will make similar updates every year or two to keep ColdFire on an upward performance track. For Motorola, ColdFire is neither its fastest nor its cheapest product line, but it does hold the broad middle ground and forms the basis of a number of ASIC and ASSP designs. As long as ColdFire keeps growing up, it will always have a parent—and customers—that continue to love it. —*J.T.*

### ■ 68K-to-ColdFire Software Translator Emerges

At long last, Motorola has produced a translator that converts assembly source code from 68000 to ColdFire. Although the

two families share a similar hardware architecture and instruction set, they are not binary compatible, and users have been forced to manually rewrite assembly code when moving from one Motorola processor to the other.

The translator, called PortASM/68K/CF, was not developed by Motorola but was licensed from a British company, MicroAPL. In addition to the basic 680x0 processors, the translator can translate code from the CPU32 and CPU32+ cores used in several 68300-series integrated processors. The translator runs under Windows 3.x, 95, NT, Solaris, and SunOS but not, ironically, on Macintosh or other Motorola-based systems. PortASM/68K/CF translates only assembly source code; binary translation remains a fond dream.

Recognizing the need for such a tool three years after the fact, Motorola distributes the translator for free; support, however, costs \$500 per year. Users can download the program from [www.motorola.com/isd](http://www.motorola.com/isd). —*J.T.*

### ■ IDT R4700 Hits 200 MHz

IDT has boosted the top speed of its FPU-equipped R4700 microprocessor to 200 MHz. The 64-bit chip is now as fast as IDT's high-end part, the R5000, and is nearly as expensive. At \$130 in large quantities, the R4700 is among the more expensive embedded processors available, although it is also one of the few with top-end floating-point capability.

Embedded 64-bit processors are few, and ones with FPUs are fewer, but QED and NEC have entered this market with lower-priced parts. The RM5270 (see MPR 10/27/97, p. 11) has an FPU as well as an L2 cache-control unit and sells for just \$100 at 200 MHz. NEC's R4310 MIPS processor (see MPR 10/27/97, p. 11), at 167 MHz, isn't quite as fast as the R4700, but it does have an FPU—and at just \$25, it's one-fifth the price of IDT's chip.

The market for FPU-equipped chips is growing quickly as high-end page printers sell in record numbers. Floating-point arithmetic is important for PostScript, and the high bandwidth of a 64-bit bus is helpful as well. IDT's target market is on a strong upward slope, but potential customers may find its prices a bit steep. —*J.T.*

### ■ New Core-Logic Support for M32R/D

Mitsubishi's unusual CPU-in-a-DRAM combination chip, the M32R/D (see MPR 5/27/96, p. 10), now has two companions. The company recently announced the M65439 and M65544, two core-logic support chips for the novel microprocessor, both of which are available immediately.

Both chips include a DRAM controller (which is not as superfluous as it sounds) for external memory, a DMA controller, interrupt logic, 16-bit timers, and at least two UARTs. The two chips differ in the PC Card and LCD interface: the '439 includes a two-slot PC Card controller, while the '544

*Continued on page 21*



# MicroJava Pushes Bytecode Performance

## Sun's MicroJava 701 Based on New Generation of PicoJava Core



by Jim Turley

At last month's Microprocessor Forum, Sun revealed details of its first microprocessor to execute Java bytecodes directly in hardware. The MicroJava 701, which isn't due until 2H98, will run at 200 MHz and deliver what Sun's Harlan McGhan believes is the best Java performance yet seen from any microprocessor. Some new design tweaks should also help the chip's performance on non-Java applications, such as C code.

No cost or price information was available, and Sun's performance figures are just estimates based on simulations, but initial results suggest that MicroJava 701 will be twice as fast as a 266-MHz Pentium II system on Java code. If Sun's initial estimates pan out and production stays on schedule, MicroJava 701 could be among the fastest, most cost-effective ways to execute Java code by late next year.

### New PicoJava 2 Core Replaces Original Design

Interestingly, Sun's chip is not based on the PicoJava core it announced at last year's Microprocessor Forum (see MPR 10/28/96, p. 28)—and which all of Sun's licensees are currently using. Instead, Sun quietly developed a newer Java core, which it now calls PicoJava 2. This design improves both Java and non-Java performance with more instruction folding and a longer pipeline that allows higher clock rates.

The new core has not been made available to Sun's Java-chip licensees (see MPR 6/17/96, p. 4). Instead, those six companies are nearing completion of their first chips based on the older PicoJava 1 design. None of the licensees has announced a schedule for these chips, but we expect the first samples to trickle out around 2Q98—the same time as Sun's 701. Even with their head start, this leaves LG, NEC, and the other PicoJava 1 licensees in an awkward position to compete with Sun. They're also about six months behind where they wanted to be when PicoJava was announced.

Sun has not licensed PicoJava 2 because the core has not been "productized." That is, the core design is in a state that only Sun's own designers can use, according to the company. Sun expects a fully portable (or exportable) version of the core to be ready in 2Q98, roughly the time the 701 and most of the original licensees' chips begin sampling.

### New Core Faster Through Fab, Pipeline Changes

The new PicoJava 2 core borrows much from its predecessor. Both cores execute about 85% of Java bytecodes directly in hardware, with the remainder trapped and emulated. Like PicoJava 1, the new core augments the bytecode instruction

set with a dozen or so "extended" instructions that allow software to manipulate caches, control registers, and absolute memory addresses—all things normally prohibited for Java programs. As before, Sun has not released the list of executed or emulated bytecodes to anyone but its licensees.

The changes between PicoJava 1 and PicoJava 2, as embodied in the MicroJava chip, are designed to improve performance. As Figure 1 shows, the pipeline has been extended to six stages from four. Instruction decoding, which used to take one cycle, is now allotted two. Likewise, the execution phase has been extended by a cycle.

In a planned 0.25-micron process, Sun expects the new core to run at up to 200 MHz, twice the target frequency for the original PicoJava 1 core in 0.35-micron technology. In the same process, we would expect PicoJava 2 to run about 33% faster than PicoJava 1.

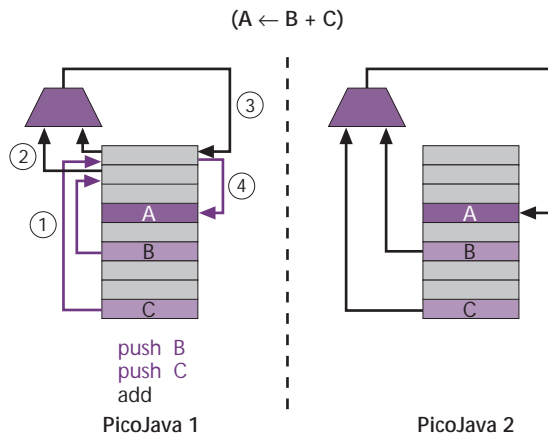
### Improved Instruction Folding Aids C Code

The other major improvement in the core design involves instruction folding. PicoJava 1 was designed to recognize certain constructs or code pairings common in Java programs (and other stack-based languages). PicoJava 2 improves on this technique by expanding the scope of the comparison. Where PicoJava 1's instruction folding assisted Java applications, PicoJava 2 should improve performance on non-Java applications as well.

Instruction folding works by recognizing certain instruction sequences that occur frequently and quietly replacing them with equivalent, but quicker, operations. In this way, the PicoJava cores address one of the bottlenecks inherent in any stack-based architecture: frequent juggling of operands on the top of the stack. Java code, for example, frequently copies one operand from the interior of the stack to the top, then uses a logical or arithmetic operation to replace the top two operands with their result. PicoJava 1 skips the preliminary copy operation and routes the first operand directly to the ALU.

Fetch	Decode	Read	Execute	Cache	Write
Fetch 1-8 bytes from cache to 16-byte instruction buffer	Decode top entries from instruction buffer	Fetch operands from stack cache; access microcode ROM	Execute; detect branches; bypass operands	Access data cache	Retire instructions; write results to stack

Figure 1. The PicoJava 2 core, which forms the basis of the MicroJava 701 chip, uses a new six-stage pipeline that breaks bytecode decoding and execution into two stages apiece.



**Figure 2.** The PicoJava 2 core is more aggressive than the original PicoJava 1 in bypassing stack manipulation. The core's decode logic recognizes common stack operations and converts them to straightforward two-input ALU functions like a conventional CPU.

The PicoJava 2 design goes one step further, routing any two operands from the interior of the stack directly to the ALU. This enhancement helps C code and other conventional languages more than it helps Java, because C compilers (and programmers) frequently use two source operands in their calculations. By fetching operands directly from the stack, PicoJava 2 more closely emulates a conventional register set, which maps much more easily onto the code engines of most compilers.

A normal C-language statement that adds two variables generally translates to a single instruction on most RISC processors. As Figure 2 shows, executing this on a stack-based architecture such as PicoJava takes from one to four instructions, depending on where the source operands happen to be in the stack and where the result will be stored. One or both source operands may have to be copied to the top of the stack before the addition can be performed, costing one

or two extra instructions. (All cases assume the operands are already loaded from memory.)

The PicoJava 2 core recognizes the case where two operands are copied to the top of the stack and replaced by their result; it then bypasses both copies, shuttles the two operands to the ALU, and stores the result in the destination register. Thus, PicoJava 2 reduces this otherwise awkward construct to a single operation, just like a RISC chip.

Note that the compiler or programmer doesn't have to be aware of the folding; PicoJava 2 does it automatically, like register renaming or instruction reordering. The object code still includes the intermediate push instructions.

### C Code Uses Back Door Into MicroJava

From a Java programmer's perspective, the PicoJava 2 core creates an infinitely deep stack. In actuality, the first 64 elements are kept in a hardware stack on the chip, while stack elements 65- $n$  spill over into external memory (or on-chip cache). From the point of view of a C compiler, PicoJava 2 has 64 general-purpose registers.

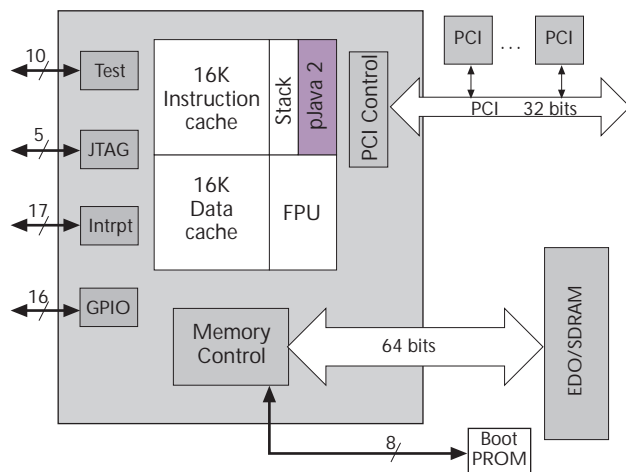
Although PicoJava 2's enhancements are designed to help non-Java code, all programs are still compiled to Java bytecodes, regardless of their original source language. Sun is developing a compiler for C and C++ that emits Java bytecode. The compiler uses the "extended" bytecodes on the 701 so programs can reference memory and access I/O devices. In perhaps the ultimate irony, such programs will not be portable because they rely on MicroJava-specific instructions rather than the nominally neutral bytecodes. It's also an incongruous reversal of the current paradigm of writing applications in Java to run on general-purpose processors.

### Tape Out Next Year; Production Not As Clear

As the block diagram in Figure 3 shows, the 701 will include dual 16K caches, a memory controller, and a 32-bit PCI interface, making the chip a nearly self-contained Java engine. The 64-bit memory controller handles EDO DRAM, SDRAM, SRAM, flash memory, and ROM. The 701 also has a separate 8-bit bus for a boot PROM. The PCI interface runs at either 33 or 66 MHz and supports both master- and target-mode operation. Overall, the capabilities of the 701 are similar to those of Sun's other embedded processor, the MicroSparc-2ep (see MPR 5/6/96, p. 5), which the company currently uses in its JavaStation 1 and JavaEngine 1 platforms.

The chip has not taped out yet, but Sun's McGhan expects the 701 to reach that milestone sometime in 1Q98, with samples in 2Q98, and full production by 3Q98. These projections may be overly optimistic; past experience has taught us that the long march from tape out to production lasts closer to 12 months, not 6. If the 701 tapes out as planned early next year, it seems likely the chip won't enter production until 1999.

Technically, the company is not behind schedule for its original claim of Java chip availability before the end of



**Figure 3.** Sun's initial MicroJava 701 implementation will include a pair of 16K caches, a 64-bit memory controller, and a PCI bus.

1997—at least, not if one of the licensees can deliver silicon early. Clearly, though, Sun has defaulted on any plans to deliver a chip of its own this year—and possibly next.

Figure 4 shows a die plot of the anticipated design. The device will measure a bit under 50 mm<sup>2</sup> and include about 2.8 million transistors, of which about 2 million go to the caches. Like most 0.25-micron microprocessors, the 701 will need dual power supplies: 2.5 V for the core, 3.3 V for I/O.

The 701 will be built in a 0.25-micron CMOS process, though the company would not identify its foundry partner. Historically, Sun has worked with Texas Instruments for most of its SPARC processors, so TI seems a likely partner.

Sun expects the chips will run at about 166 MHz, with a useful percentage yielding at 200 MHz. At the faster speed, the 701 should consume about 4 W, according to company estimates, and the chip will come packaged in a 316-contact plastic ball-grid array (PBGA). The MDR Cost Model yields an estimated manufacturing cost of \$25 for the part.

### Java Performance Beats Pentium II by 2×

Sun has been unusually tight-lipped about the performance of its Java chips. More than a year after announcing PicoJava 1, the company still has no verifiable performance metrics.

At the Forum, McGhan revealed that Sun has simulated the 701 running the CaffeineMark and Dhrystone benchmarks. The tests yielded a rating of 200 MIPS on Dhrystone 2.1 and 13,332 on Embedded CaffeineMark 3.0. The Dhrystone score is a little below average for a 200-MHz chip; the Embedded CaffeineMark score, however, is far higher than anything seen before.

Specifically, the highest rating recognized by Pendragon Software (the creator of CaffeineMark; [www.webfayre.com](http://www.webfayre.com)) is 7,379 for a 266-MHz Pentium II system with 64M of RAM running Windows NT and Internet Explorer. (Embedded CaffeineMark eliminates three of the nine tests from the full CaffeineMark 3.0 suite, for a higher overall rating.)

Thus, Sun's simulations indicate the 701 executes Java almost twice as fast as the high-end Intel system, even though Pentium II has superscalar execution, a one-third faster clock rate, and larger caches. It's also an order of magnitude faster than StrongArm. At 233 MHz, the SA-110 scores just 1,105 on Embedded CaffeineMark 3.0, a disappointing 12× slower than the 701, even at a slightly faster clock speed.

### Rockwell Still a Wildcard

MicroJava will also be up against JEM1, Rockwell's surprise entry to the Java field (see MPR 10/27/97, p. 10). The core of this come-from-behind player, which is based on an old avionics processor from the company's archives, is smaller than PicoJava and executes more bytecodes in hardware.

At just 50–60 MHz in 0.5-micron technology, JEM1 is not as fast as the 701. It would speed up considerably, though, if it were built in the same 0.25-micron process. Rockwell and Sun are said to be negotiating a distribution agreement for JEM1; it would be interesting if some of

JEM1's design features appear in a future MicroJava processor. Rockwell has no benchmark information whatsoever for JEM1, and no price has been set, so the benefits of this chip are impossible to judge.

### Dhrystone Performance Not As Good

To the extent that one trusts Dhrystone, the SA-110 and Pentium II both do much better than the MicroJava chip. The StrongArm chip rates at 268 Dhrystone MIPS, versus 200 MIPS for MicroJava. Pentium II scores range from 300 to 400 on Dhrystone at 233 MHz, putting it 1.5× to 2× ahead of the 701. Both results are measured on real systems.

On the other hand, Sun's results are simulated, and these benchmarks are too small to accurately reflect cache misses or memory latency. When the 701 begins shipping, its actual score may be different. At the same time, we can assume JIT compilers and other microprocessors will only get faster. By 3Q98, Pentium II should be shipping at 400 MHz (at least) in the same 0.25-micron process as the 701.

Still, Sun's results are impressive, even as a first-order approximation. To deliver performance in the same range as a Pentium II—much less beat it by 2×—with a chip that's half the size and (presumably) less expensive is no small feat. Factoring in the memory savings (because the 701 replaces a full Java interpreter or JIT compiler with a small emulation library), MicroJava 701 looks to be a dynamite bargain for customers determined to build Java-execution machines.

### Waiting for the Demand

The question, of course, is exactly what kind of machines those might be. The hypothetical Java-based network computer has been slow to appear, perhaps because useful Java applications are not thick on the ground. Corel, for example,

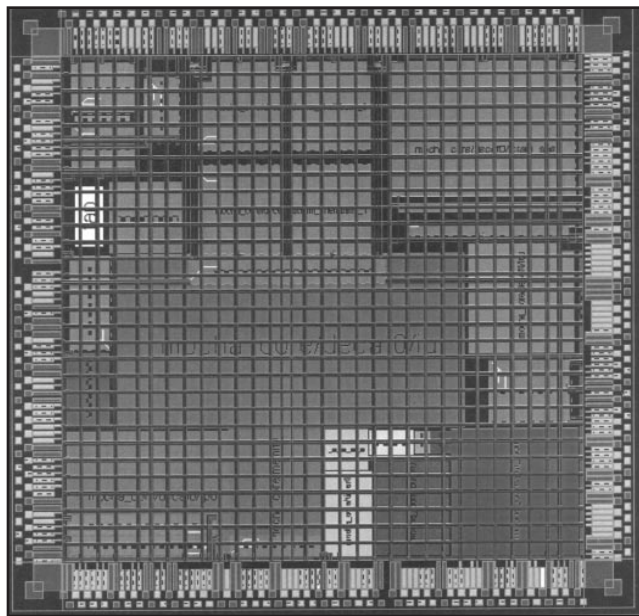


Figure 4. Die plot of proposed MicroJava 701 layout indicates the chip will measure about 50 mm<sup>2</sup> in a 0.25-micron CMOS process.



## Price & Availability

Samples of Sun's MicroJava 701 are expected to be available in 2Q98, with production in 2H98. Pricing has not been announced. For more information, contact Sun Microelectronics (Mountain View, Calif.) at 650.960.1300 or visit [www.sun.com/microelectronics/java](http://www.sun.com/microelectronics/java).

canceled its high-profile attempt to port its WordPerfect Suite to Java. Without plentiful Java apps, Java systems are superfluous; without the Java systems, the apps may not come, so perhaps Java NCs are just an egg waiting for a chicken.

Even assuming demand for such a system, a Java chip is just one of many options. A general-purpose microprocessor leaves the door open to other languages, APIs, and operating systems besides Java, Java, and Java. By the time the 701 appears, it may be no faster than Pentium II on Java code, but with an infinitely smaller software base. Whereas general-purpose processors can execute anything the 701 can, the reverse condition does not hold true.

Although McGhan would quantify the expected selling price of the initial MicroJava chip only as "two digits," it's safe to assume that the 701 will be much less expensive than Pentium II, thus providing a price/performance advantage to developers for whom software availability is not important. But the same price/performance claim could be made of most other microprocessors as well.

### Java Bytecode Sets Strategy

With the impending arrival of Java chips, embedded-software developers will soon be faced with three basic alternatives: write in C; write in Java; or write in Java and compile to bytecodes. In at least two of the three cases, Java chips do not make a compelling argument.

For the C-to-native scenario, the 701 makes very little sense. C programs written for the 701 are no more portable than other compiled programs, and the 701's performance (if Dhrystone is any indication) isn't particularly good.

The Java-to-native scenario also favors general-purpose microprocessors. Compiling Java source directly to the native instruction set of the target microprocessor bypasses the bytecode interface, skipping a costly intermediate step. Bytecode was intended for portability; if the software isn't being ported, it serves no purpose. This approach may sacrifice the putative portability of bytecode, but for embedded systems, real-time binary portability is rarely an issue.

Finally, there is the Java-to-bytecode scenario. If bytecode is the preferred delivery mechanism, the 701 will run it

quickly and with minimal memory overhead. But in return, the chip exacts a toll in the use of non-Java software, operating systems, tools, and APIs.

Nearly any system can download and run the occasional Java applet. The advantage of the 701 is running those downloaded applets *quickly*. For "casual" use of bytecode, where performance is not all-important, a general-purpose processor can handle the task, and give better overall performance when it's not running bytecode.

In short, the 701 looks better the more bytecode the system has to run. For an all-bytecode system, the 701 is probably faster and cheaper than anything else. As the proportion of bytecode decreases, so does the advantage of a dedicated Java chip. MicroJava 701 and its kind make sense for some small fraction of the market (that does not now exist) that mainly relies on Java code and doesn't already have a microprocessor in it.


### Java: Doing Whatever It Takes

It's no secret that Sun has focused its corporate efforts on the success of Java. Java hardware, software, education, and advertising are the company's featured products.

Strategically, Sun is more interested in Java itself than in Java chips specifically. McGhan was careful to point out that Java chips wouldn't and shouldn't replace Java interpreters or JIT compilers, but that they merely bring another option to the table. Java chips are "a complement, not a replacement" for software-only Java environments, he avowed.

At the level of the executive suite, Sun doesn't really care whether Java chips succeed or fail. Sun's ultimate goal is that Java prevail, through whatever means. The company is offering as many different methods of writing, disseminating, and executing

Java code as it knows how. Whether customers execute Java applications using interpreters, JIT compilers, or specialized Java chips is irrelevant, as long as they use Java instead of Microsoft APIs.

Like a heavy shovel, Java has left a lasting impression on the minds of designers of both desktop and embedded systems. As companies wrestle with questions about whether they want Java, where to use Java, and how to execute Java, Sun has fanned the flames and encouraged experimentation. For the time being, the experimenters and the tire kickers have been using general-purpose microprocessors with Java interpreters, Java compilers, and Java-aware operating systems. For another 6-9 months, this will probably still be the case. Not until MicroJava 701, JEM1, or one of Sun's licensees' parts starts shipping will Java's early adopters be able to see for themselves whether a dedicated Java processor is valuable for their application. 



MICHAEL MUSTACCHI

Harlan McGhan of Sun Microelectronics extols the virtues of the MicroJava 701 at the Forum.

# Siemens TriCore Revives CISC Techniques

## New 32-Bit Design Emphasizes DSP Capability and Microcontroller Functions



by Jim Turley

Fashions run in cycles, and CPU designs abandoned in the '70s are again becoming chic. As evidence of this trend, Siemens rolled out TriCore, a retro-CISC architecture that combines microcontroller, DSP, and CPU features while simultaneously attempting to shrink code size and improve speed.

TriCore throws out nearly everything that RISC microprocessor design has taught in the past ten years. At last month's Microprocessor Forum, TriCore's chief architect, Rod Fleck, described a mixture of 16- and 32-bit instructions, separate address and data registers, multicycle instructions, multiple data types, complex interrupt handling, and a lopsided instruction set geared toward bit-twiddling.

Paradoxically, TriCore also includes some of today's newest thinking in embedded controllers. It handles digital signal processing (DSP), zero-overhead loops, and SIMD packed data types, and relies on a 128-bit path to local embedded DRAM. The first TriCore chips aren't expected to sample until 2Q98, but by the end of next year, Siemens hopes to make a dent in automotive and computer-peripheral markets with its unusual new family.

### Register Set a Big Step Backward

TriCore includes 32 registers, plus a few control/status registers. Where TriCore differs from most recent microprocessors is its split between address and data registers. As Figure 1 shows, only 16 of the registers are used for handling data; the other half are address pointers.

TriCore's register set is further divided into quarters. Siemens segregates the registers into an "upper context" and a "lower context." When switching tasks, calling functions, or handling interrupts, only one context is saved or restored.

TriCore's lower context consists of eight data registers (D0-D7) but only six of the lower eight address registers (A2-A7). The other two address registers, A0 and A1, are global resources that are not saved or restored as part of a context switch. The global program counter, PC, is also considered part of the lower context. The upper context consists of the other eight data registers (D8-D15), six address registers (A10-A15), and PSW. Like A0 and A1, A8 and A9 are global address pointers.

TriCore's subdivided (and not very general) register set is a throwback to the CISC microcontrollers that Siemens knows so well. While the design might not be architecturally elegant, Fleck believes that it makes sense in the grubby world of real-life control code. The logical split allows

Siemens to implement a physical split, placing the addresses and data registers in different areas of the chip, alleviating port congestion and simplifying routing.

### Linked List Gives Fast Interrupt Response

Interrupts, traps, and function calls automatically save the upper context to on-chip memory, giving the trap handler, interrupt-service routine, or called function a clean set of registers with which to work. Software can save the lower context as well, if desired, with explicit instructions. TriCore automatically restores the upper context when resuming the interrupted task. Saving and restoring the upper or lower context takes just four clock cycles because of TriCore's wide 128-bit bus between the register file and dual 2K caches.

Saved contexts are stored in a linked list that TriCore chips maintain automatically. Each context store needs 64 bytes for the 15 registers plus a 32-bit link pointer to the next free context area. Internal head and tail pointers (in PCXI) allow TriCore to quickly locate the next (or previous) state.

As a safety net, TriCore also maintains a call-recursion counter for every task. The counter is incremented on every function call and restored to its previous state (rather than decremented) on every return. If the counter overflows, an exception occurs. Programmers can control the width of the counter (in bits), setting their tolerance factor for runaway recursion.

### Instructions an Eclectic Mix

TriCore executes both 16- and 32-bit instructions, which may be freely intermixed. Each opcode includes a size bit, so TriCore's instruction decoder can identify long or short instructions immediately. The 16-bit instructions form a

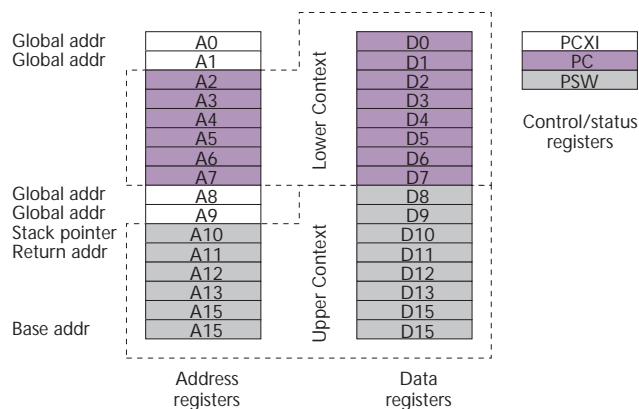


Figure 1. TriCore separates its address and data registers; upper and lower halves, or contexts, are saved and restored separately.

subset of existing 32-bit instructions that sacrifices some flexibility, so in many cases a 32-bit instruction can be replaced with a shorter form to save code space. Unlike Thumb or MIPS-16, TriCore switches between 16- and 32-bit instructions without explicitly switching modes, saving a little in execution time and avoiding the need to segregate compressed and uncompressed code.

TriCore follows a basic load/store model, but the similarity to RISC chips ends there. The instruction set is quite rich, with most of the emphasis on numerical processing and integer DSP work. As Table 1 shows, TriCore has seven different forms of addition, for example, and a dozen different multiply-accumulate instructions. Far from being general purpose, TriCore was developed for motion control, signal processing, and compression/decompression work.

Most arithmetic operations can operate on bytes, halfwords (16-bit quantities in the Siemens argot), words, doublewords, and so-called Q-format numbers. Parallel (SIMD) byte and halfword operations are richly supported.

The Q15 and Q31 data types are simplified fractional representations, with one sign bit followed by 15 (or 31) bits of significance after the implied binary point. The advantage of Q-format over IEEE-754 floating point is that it is far easier to implement in silicon. Left-justified Q15 and Q31 numbers can be added by normal addition instructions without compromising precision. On the other hand, Q-format numbers do not have the dynamic range or flexible precision of true floating-point numbers. For many control systems, Q-format will do the trick with low hardware cost. For applications that require true floating-point math, however, Q-format won't cut it.

### Strong Parallels and Bit Wise

In addition to the usual operations, TriCore can perform integer or Q-format arithmetic on packed data values. In a manner similar to MMX, two halfword values or four byte values can be added, subtracted, multiplied, multiply-added, or multiply-subtracted. The ADD.B instruction, for example,

Mnemonic	Description	Mnemonic	Description	Mnemonic	Description
<b>Arithmetic</b>		<b>Load/Store</b>		<b>Multiply/Divide</b>	
ADD(S)	Add (with saturation)	MOV	Copy register	MUL	Multiply 32 × 32 → 32
ADDC	Add with carry	LD	Load	MULS/R	... with saturation/rounding
ADDI	Add immediate value	ST	Store	MULM	Multiply 32 × 32 → 64
ADDIH	Add immediate to high half	LEA	Load effective address	MADD	Multiply-add 32 bits
ADDSC	Add scaled value	LDLCX	Load lower context	MADDS/R	... with saturation/rounding
ADDX	Add and generate carry	LDUCX	Load upper context	MADDRS	... with saturation and rounding
SUB(S)	Subtract (with saturation)	STLCX	Store lower context	MADDM(S)	Multiply-add 64 bits (with sat.)
SUBC	Subtract with borrow	STUCX	Store upper context	MSUB	Multiply-subtract 32 bits
SUBX	Subtract, generate borrow	SVLCX	Save lower context	MSUBS/R	... with saturation/rounding
RSUB(S)	Reverse subtract (with saturation)	RSLCX	Restore lower context	MSUBRS	... with saturation and rounding
ABS(S)	Absolute value (with saturation)	LDMDST	Load, modify, store	MSUBM	Multiply-subtract 64 bits
ABSDIF(S)	Absolute value difference (sat.)	MFCR	Move from special register	MSUBMS	... with saturation
DIFSC	Difference scaled addresses	MTCR	Move to special register	DVADJ	Adjust after division
CLO/CLZ	Count leading ones/zeros	SWAPA	Swap with address register	DVINIT	Prepare for division
CLS	Count leading signs	<b>Comparison</b>		DVSTEP	Division step
MIN/MAX	Find minimum/maximum	EQ	Compare, equal	<b>Bit Manipulation</b>	
SAT	Saturate operand	NE	Compare, not equal	DEXTR	Extract from doubleword
<b>Logical</b>		GE	Compare, greater equal	IMASK	Create mask word
AND/OR	Logical AND/OR	LT	Compare, less than	INS	Insert single bit
NAND	Logical NAND	EQANY	Compare for equality	INSERT	Insert bit field
NOR	Logical NOR	EQZ	Compare address for zero	EXTR	Extract bit field
XOR	Logical exclusive-OR	<b>System</b>		<b>Flow Control</b>	
XNOR	Logical exclusive-NOR	ENABLE	Enable interrupts	J	Jump, unconditional
ANDN	Logical AND 1's comp.	DISABLE	Disable interrupts	JA/JI	Jump, absolute/indirect address
ORN	Logical OR 1's complement	DSYNC	Force pending data accesses	Jcc	Jump on condition cc
NOT	Logical invert	ISYNC	Flush execution pipeline	JL	Jump and link
SH/SHA	Logica/arithmetic shift	RSTV	Reset overflow flags	JLA/JLI	... absolute/indirect
SHAS	Arithmetic shift with saturation	SYSCALL	Force trap	LOOP	Initiate loop
<b>Conditional</b>		TRAPV/SV	Trap on overflow	BISR	Begin ISR
CADD(N)	Conditional add (1's comp.)	DEBUG	Enter debug mode	CALL	Call subroutine
CSUB(N)	Conditional subtract (1's comp.)	NOP	No operation	CALLA/I	Call absolute/indirect
CMOV(N)	Conditional move (1's comp.)			RET	Return from subroutine
SEL	Select operand		16-bit instruction word	RFE	Return from exception

Table 1. TriCore has an unusually rich instruction set, including several 16-bit instructions (shaded). TriCore's designers paid particular attention to multiplication, accumulation, and reverse addition and subtraction—all functions used in digital-signal processing.



adds the four bytes from two registers and deposits four results in another register. As with most arithmetic operations, saturating, nonsaturating, signed, and unsigned variations are all available.

Parallel logical operations and comparisons are also supported. The EQANY.B instruction, for example, compares each byte from two registers and sets the LSB in the destination register if any of the four comparisons is equal. An alternate form compares all four bytes against a constant.

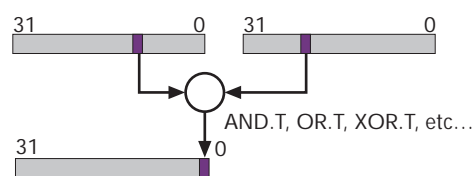
Moving beyond simple arithmetic, TriCore implements a bewildering combination of multiply, multiply-add, and multiply-subtract instructions with options for saturation, rounding, sign extension, packed data, integer or Q-format data types, alignment, and addressing modes. These, plus the zero-overhead LOOP, form the heart of TriCore's DSP capability. Combined with bit-reverse and circular memory addressing, TriCore will become a creditable digital-signal processor, Siemens believes. Simulations show the chip can sustain an FIR filter at two taps per cycle, twice the throughput of either Motorola's 56300 or TI's popular 'C54x parts.

Well beyond the ken of most microprocessors are TriCore's bit-manipulation instructions, such as INSERT, INS, and EXTR. The first two insert a single bit, or a set of contiguous bits, into any location in another register; the last instruction extracts any number of bits from any location in any register and copies them into another register, with either sign- or zero-extension to 32 bits. For I/O control and network addressing, among other applications, these operations are valuable.

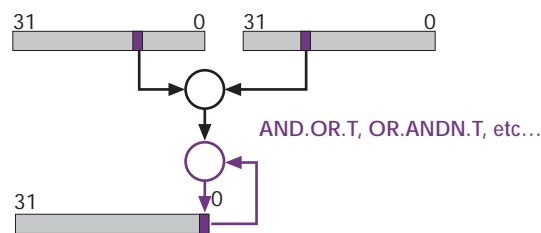
The CLO, CLZ, and CLS instructions count leading ones, zeros, or sign bits, respectively. The variants CLx.H and CLx.B are interesting because they count leading bits in packed bytes or halfwords, returning two or four totals. All of these operations are useful for normalization, prioritization, encryption, error correction, and for managing graphics primitives, and they are time-consuming to perform any other way.

Conditional operations abound, even though TriCore has no condition codes. Conditional adds, subtracts, moves, and branches all depend on comparing a register (always D15 in 16-bit instructions) with zero. Conditional moves CMOV and CMOVN, for example, copy the contents of a register only if D15 is zero or nonzero, respectively. SEL is like CMOV, but it copies one of two values, depending on D15. Other conditional instructions can test any data register.

Integer division comes in the form of DVINIT, DVSTEP, and DVADJ, three instructions that perform step-wise division under software control. TriCore's support for division is better than in many older CPU families that have no divide instruction at all, but for ease of use it falls short of the single-instruction divide operations in the 68K and x86. At one bit per cycle, TriCore is the same speed as SuperH, which also requires explicit divide-step instructions. TriCore's DVSTEP always works in eight-cycle chunks, though.



(a) Normal two-input Boolean operation



(b) Accumulating, three-input Boolean operation

Figure 2. TriCore's normal's two-input logical operations are complemented with a set of eight accumulating, three-input logicals.

### Three-Way Logicals Tighten Code

One of the most interesting and unusual features of TriCore's instruction set is its ability to "accumulate" the results of logical (Boolean) operations. Using more than a dozen instructions for accumulating and three-way comparisons, programmers can implement complex multiway comparisons in a minimum amount of code.

As Figure 2a shows, the normal bit-wise comparison instructions, such as AND.T, OR.T, or XNOR.T, compare any two bits from any two registers and store the logical result in the least-significant bit of a third register (the rest of the destination register is cleared). All four basic Boolean operations (AND, OR, XOR, ANDN) and their opposites (NAND, NOR, XNOR, ORN) are available. Table 2 lists the accumulating logical operations.

TriCore goes a step further by allowing three-operand logicals. As Figure 2b shows, the result of a previous bit-wise logical operation can be included in a new logical operation, like a multiply-accumulate on a single bit. The logical opera-

Logical			
AND.AND.T	OR.AND.T	SH.AND.T	SH.NAND.T
AND.ANDN.T	OR.ANDN.T	SH.ANDN.T	SH.ORN.T
AND.OR.T	OR.OR.T	SH.OR.T	SH.XOR.T
AND.NOR.T	OR.NOR.T	SH.NOR.T	SH.XNOR.T
Comparison			
OR.EQ	AND.EQ	XOR.EQ	
OR.NE	AND.NE	XOR.NE	
OR.GE	AND.GE	XOR.GE	
OR.GE.U	AND.GE.U	XOR.GE.U	
OR.LT	AND.LT	XOR.LT	
OR.LT.U	AND.LT.U	XOR.LT.U	

Table 2. TriCore can perform two parallel logical operations or arithmetic comparisons simultaneously. The Boolean result can be "accumulated" with other results, allowing programmers to efficiently code sequences of comparisons and/or logical operations.

## Price & Availability

The first Siemens TriCore chip (which has not been named) is expected to sample in 2Q98; production is set for 2H98. For more information, contact Siemens Microelectronics (Cupertino, Calif.) at 408.895.5004 or visit [www.sci.siemens.com/tricore](http://www.sci.siemens.com/tricore).

tions don't even have to be the same, so the first and second bits can be AND-ed while the third bit is OR-ed, for example. This construct maps well to common multipart comparisons in high-level languages and allows programmers (or compilers) to efficiently encode complex relational tests.

The comparisons need not be limited to single bits, either. Using instructions like OR.EQ, AND.LT, or XOR.GE, programs can cumulatively compare 32-bit numbers using signed or unsigned arithmetic comparisons. By accumulating a running Boolean result through multiple comparisons, TriCore programmers can avoid peppering their code with a lot of short conditional branches that bloat code size and lead to pipeline bubbles.

### Instruction Set Hardly Orthogonal

TriCore is not without its peculiarities. Its split register set will generate split opinions; separate address and data registers are shunned these days. The argument follows that for split versus unified caches: some algorithms use more data, while others need more address pointers. Unifying the register file (or the cache) allows dynamic allocation of a scarce resource.

The split register file also means separate instructions for the address registers, crowding TriCore's opcode map. Instructions like ADD.A and SUB.A are needed to update address pointers; special MOV instructions move the contents of address registers around. A split in the data path also forces TriCore to keep its loop counter in an address register to relieve congestion of the data registers in filter operations.

Flow-control instructions with absolute addresses are limited to 24-bit pointers, at best. TriCore handles this by separating the four most significant bits and left-justifying them, then using the lower address bits unchanged. This technique sections TriCore's 4G address space into 16 equal-sized regions, with the lower 1M of each directly accessible to other code.

### Ugly Has a Place, Too

With TriCore, Siemens has thrown out the book of accepted design principles and relied on its years of experience designing and selling microcontrollers to industrial and

commercial OEMs. TriCore is different, unconventional, and sometimes awkward, but it seems to have the tools to get the job done quickly.

TriCore is only the second microprocessor architecture to be designed from the ground up around embedded DRAM (Mitsubishi's M32R/D is the other). TriCore's context-switch mechanism relies on a fast, wide connection to on-chip memory, and Siemens's Fleck says no TriCore chip will be made without it.

TriCore is also one of the few microprocessors to deal with control and signal processing in a single instruction set. ARM's Piccolo and Hitachi's SH-DSP were both grafted onto the original design and force tradeoffs in DSP-versus-controller performance. But those grafts can also be rejected if users are interested in just the core ARM or SuperH processor, a choice that Siemens doesn't offer.

Motorola's ColdFire and M•Core lines have both taken their first small steps toward signal processing, but no 32-bit processor out there can match the dizzying assortment of number- and bit-manipulation options TriCore will have. The important details, like bit-reverse and circular addressing, rounding, saturation, and sticky overflow bits, all give TriCore a much more credible claim to being a DSP than most microprocessors can make. On the downside, TriCore doesn't offer the X/Y data memories DSP programmers admire, but that don't map well to C code.

The inevitable price for this complexity is circuit density. Siemens has not released any details of die size, transistor count, or clock speed. The first TriCore chip, with 128K of on-chip DRAM and a pair of 2K caches, has not taped out and samples aren't expected until 2Q98. That's a

long time to wait before customers can make informed choices—time that Motorola, Hitachi, and untold hordes of ARM vendors can use to strengthen their leads.

There's a reason CPU designers have scrapped complex and irregular instruction sets and embraced—more or less—the principles of RISC design in recent years. Simpler architectures make simpler CPU cores, which run faster and are easier to scale to multiple execution units. During his presentation, Siemens's Fleck hinted at plans for TriCore chips with more execution units in the not-too-distant future, but with all of TriCore's complexity, that may be tough to do.

With production volumes probably coming in 1999, it's still too early to be certain about much to do with TriCore. Certainly the family will compete with ARM, ColdFire, M•Core, SuperH, and MIPS chips. But for those applications that emphasize signal processing over numerical processing, bit-tiddling over floating-point handling, and code density over software compatibility, TriCore appears well positioned to pick up a portion of new embedded designs.  $\square$



MICHAEL MUSTACCHI

**Siemens CPU director Rod Fleck talks about the CPU/DSP capabilities of TriCore at the Forum.**

# Centaur Improves C6 With No Extra Cost

*New C6+ Is Faster on Integer, FP, and MMX Applications; Original C6 Ships*



by Linley Gwennap

Even as it begins selling its initial x86 processor, the C6, IDT subsidiary Centaur has already developed a follow-on product with many minor improvements. The C6+, due to reach production in 2Q98, brings MMX and FP performance up to par with the competition while adding other features, such as branch prediction, that raise performance on general PC applications. Centaur remains focused on low cost: these improvements add just 3 mm<sup>2</sup> to the die size of the C6.

Speaking at last month's Microprocessor Forum, Centaur founder Glenn Henry disclosed plans to further boost the performance of the C6, now known as the IDT WinChip family. IC process improvements should allow the part to reach 300 MHz by mid-1998. In 2H98, IDT will use a 0.25-micron process to further increase clock speed while incorporating a 256K on-chip level-two cache with the CPU.

Centaur has independently adopted many of the same techniques being used by AMD and Cyrix to extend their K6 and 6x86MX processors, respectively, including an enhanced FPU, new instructions for 3D graphics, a faster system bus, and an integrated L2 cache. Although the C6 has a die-size advantage over competitive 0.35-micron processors, AMD plans to begin shipping its 0.25-micron K6, with a smaller die than the C6, in 1H98 (see MPR 10/27/97, p. 19).

The C6+ tapeout is imminent but has not yet occurred, so all performance projections are based on simulations. Because of the limited number of changes from the current

C6 and the team's demonstrated skills, Henry believes the new part will reach volume production just six months after tapeout, half the time it usually takes for a new processor.

## FP and MMX Match Pentium/MMX

The original C6 was designed and put into production in less than two and a half years, a task that takes most vendors three or four years. Given the tight schedule, the designers focused on time to market rather than on maximum performance. One area that didn't get as much attention was the floating-point unit, since the number of PC applications that use FP today is still relatively small.

The C6+ contains a completely redesigned FPU that fits into the same die area as the current unit but is now fully pipelined for most FP operations. The C6+ takes an extra cycle for double-precision multiplies, but this operation is typically used only in technical applications (e.g., CAD) that are not in Centaur's target market.

As Table 1 shows, on most instructions the new FPU is as fast as Pentium/MMX's, and Centaur expects the C6+ to deliver about the same performance on most FP applications as Pentium/MMX at the same clock speed.

Centaur also completely redesigned the C6's MMX unit to improve performance. The C6+ can issue and execute up to two MMX instructions per cycle, with essentially the same restrictions as Pentium/MMX. Except for the MMX unit, however, the C6+ is a scalar processor, so it can't pair an MMX instruction with an integer instruction.

The Centaur chip is faster than Pentium/MMX on several MMX operations. As Table 1 shows, the C6+ has better latency on MMX multiply or multiply-add, but most applications depend on throughput, not latency, for performance. Stores take a single cycle, half as long as on the Intel chip. As a result, the C6+ should slightly outperform Pentium/MMX on most MMX-based benchmarks, according to Centaur.

## New Instructions Aid 3D Graphics

Like AMD and Cyrix (see MPR 10/27/97, p. 22), Centaur has added new instructions to its next part to speed 3D geometry calculations. Centaur has taken a more radical approach, however, that is not compatible with either the AMD 3D instructions or Cyrix's MMXFP. Instead of simply increasing throughput by pairing single-precision FP values and operating on them in parallel, the C6+ implements a new set of FP registers and instructions that use them.

The new part adds 22 directly addressable 80-bit floating-point registers to the eight-entry FP stack defined by the x86 architecture. This greatly increases the number of values that can be stored in the chip, reducing time-wasting memory

	Centaur C6+		Pentium/MMX	
	Thruput	Latency	Thruput	Latency
FP add	1 cycle	3 cycles	1 cycle	3 cycles
FP store	1 cycle	n/a	2 cycles	n/a
FP multiply (SP)	1 cycle	3 cycles	1 cycle	3 cycles
FP multiply (DP)	2 cycles	4 cycles	1 cycle	3 cycles
FP mul/add*	1 cycle	3 cycles	4 cycles	6 cycles
FP to integer*	1 cycle	3 cycles	6 cycles	6 cycles
FP sq root (SP)*	24 cycles	24 cycles	70 cycles	70 cycles
FP inv sq root (SP)*	24 cycles	24 cycles	109 cycles	109 cycles
MMX add	1 cycle	1 cycle	1 cycle	1 cycle
MMX multiply	1 cycle	1 cycle	1 cycle	3 cycles
MMX mul/add	1 cycle	1 cycle	1 cycle	3 cycles
MMX store	1 cycle	n/a	2 cycles	n/a

Table 1. Compared with the Intel Pentium/MMX, the C6+ has the same or better performance on most MMX and FP instructions except for double-precision FP multiplication. \*C6+ uses proprietary instructions; Pentium/MMX uses standard x86 instructions. n/a=not applicable (Source: vendors)



## Price and Availability

IDT is now shipping its Winchip (C6) at clock speeds of 180 and 200 MHz. The "suggested retail price" is \$90 and \$135, respectively. The company did not reveal its 1,000-piece price, which is presumably lower. IDT expects to sample 225- and 240-MHz parts this month, with production shipments in 1Q98. The C6+ is scheduled to ship in 2Q98. For more information, try [www.winchip.com](http://www.winchip.com).

accesses. The designers then added 53 new instructions (using 12 x86 opcodes) to operate on these new registers. These are fully IEEE compliant and handle all precisions.

In addition to the usual arithmetic and load/store operations, the new instructions include a fully pipelined floating-point multiply-accumulate. The x86 instruction set has no such operation, and issuing a multiply and a dependent add takes four cycles on a Pentium/MMX. The new instructions also include fast square root and inverse square root operations, as Table 1 shows. These operations are frequently used in lighting calculations and other 3D geometry algorithms.

### Speed Gains Require Direct3D

As a result, Henry claims the C6+ will have significantly better performance on 3D games than Pentium/MMX at the same clock speed. Yet the die size impact of the new instructions is less than 1 mm<sup>2</sup>, mostly for the larger register file. Decoding the new opcodes is simple, and most of the data paths were already in place, supporting microcode primitives to execute existing x86 instructions.

These improvements are moot, however, unless the new instructions are used by software, a struggle with which Centaur, along with AMD and Cyrix, must contend (see MPR 10/27/97, p. 35). Centaur is developing its own version of Microsoft's immediate- and retained-mode library for Direct3D. If Microsoft agrees to distribute this code as part of DirectX 6 and Windows 98, any 3D application that uses this API could take advantage of the new instructions transparently.

Microsoft has supplied its code to Centaur but hasn't committed to distributing Centaur's version. Instead, Microsoft would prefer that the x86 vendors agree on a single set of 3D extensions, and Henry has volunteered to adapt his chip to one of the others. At the Microprocessor Forum, AMD CEO Jerry Sanders publicly offered to license his company's AMD 3D extensions, and the K6 3D will ship before Cyrix's Cayenne, making AMD the logical choice. Henry couldn't confirm whether he will use AMD's or Cyrix's extensions.

Centaur has gone a step further than any other vendor, even Intel, by adding state to the machine. Unless Microsoft agrees to modify Windows 98, which seems unlikely, the state of the 22 new registers will not be saved and restored on a context switch. If the new instructions are used only for 3D, and only one 3D application is running at a time, there should be no data corruption.

A further concern is the potential incompatibility with future instructions. If Intel were to use one of the 12 new opcodes for a different purpose, say an instruction in its forthcoming MMX2 extensions, it would be difficult for Centaur to implement both its instructions and the new Intel instructions. Unless Intel reveals its MMX2 encodings soon, it could put Centaur over a barrel.

### Standard PC Applications Gain About 6%

The C6+ includes several minor improvements to speed standard integer applications. Some couldn't be implemented in the C6 due to its tight schedule; others became apparent later when analyzing code traces for performance bottlenecks.

Taken together, these changes improve performance over the C6 by about 6% on the Winstone 97 Business benchmark, a collection of popular PC applications.

For example, the C6+ improves the timing of several instructions. Integer multiplies take 6 cycles (compared with 10 for Pentium/MMX). Arithmetic instructions that write their results to memory take two cycles, one fewer than on the Intel chip. The otherwise scalar execution core can pair two PUSH or two POP instructions, emulating the two-way superscalar Pentium core.

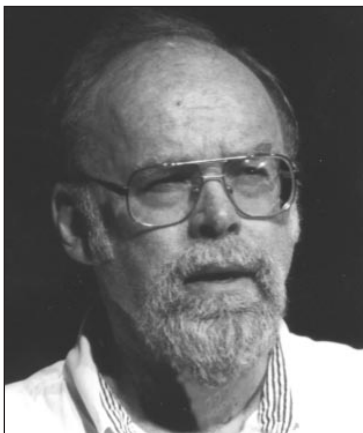
Cache improvements also contribute to better integer performance. The new data cache is four-way set associative, twice the number of sets as in the C6. The new cache also supports a write-allocate option,

reducing the number of writes to the L2 cache when this mode is enabled.

### Highly Accurate Branch Prediction

Perhaps the most important change is adding branch prediction. The C6, like the 486, has no branch prediction and takes a three-cycle penalty on every taken branch (except subroutine returns, which are handled by a return-address stack). The C6+ has a 4,096 × 1-bit branch history table (BHT) that predicts the direction of each branch but not the target address, as Pentium's branch target buffer (BTB) does.

Thus, correctly predicted taken branches create a single-cycle "bubble" in the execution stream. Because the C6 uses an instruction queue to decouple the fetch stream from the execution units (see MPR 6/2/97, p. 1), this bubble reduces performance only when the queue is empty, which is about 15% of the time. By not caching target addresses, the C6+ is



MICHAEL MUSTACCHI

Centaur founder Glenn Henry describes the new features of the C6+ at the Microprocessor Forum.

able to track the history of eight times as many branches as Pentium/MMX, improving its prediction accuracy.

Accuracy is further improved by two advances in the prediction algorithm. The C6+ uses a two-level indexing method known as Gshare and a new encoding method called “agrees” (see page 22); Intel has not provided the details of Pentium/MMX’s two-level BTB. As a result of these changes and the larger BHT, the C6+ correctly predicts 93% of the branches in Winstone 97 Business, compared with 82% for Pentium/MMX, according to Centaur. Yet the C6+ BHT consumes less than one-eighth of the space of Intel’s BTB.

### Roadmap Includes On-Chip L2 Cache

IDT is currently selling the C6, which is built in a 0.28-micron hybrid process (the metal layers are similar to those of a 0.35-micron process), at speeds of 180 and 200 MHz. The part is currently sampling at speeds of 225 and 240 MHz, and the company plans to ship these speed grades in 1Q98.

Since the C6+ is essentially the same size as its predecessor, IDT plans to offer it at the same price, providing more performance for free. In addition, circuit rework should boost its performance to 266 MHz in the same process. IDT also has a version of the 0.28-micron process that reduces the supply voltage from 3.3 V to 2.5 V; this reduction allows thinner gate oxides and thus faster transistors. In this process, the C6+ is projected to reach 300 MHz by mid-1998, as Figure 1 shows.

In 2H98, IDT plans to shrink the C6+ to a true 0.25-micron process. This process should reduce the die size to less than 60 mm<sup>2</sup>, smaller than any competitive x86 processor. Running at 2.5 V, the chip should have moderate power dissipation despite higher clock speeds, making this new part suitable for notebooks or desktops. Regarding clock speed, Henry says only that the 0.25-micron C6+ should operate in excess of 300 MHz.

To take advantage of the tiny CPU core as well as IDT’s experience as a leading SRAM vendor, the company plans to pack a 256K level-two cache onto the same die as the CPU, providing an integrated product along with the nonintegrated version. This cache will be eight-way set associative and operate at the full processor speed, offering much better performance than an external cache. In fact, this design provides many of the benefits of Intel’s Pentium II dual-bus architecture without breaking compatibility with Socket 7. The integrated part will be particularly good for notebooks, where the power savings should be substantial.

This part appears to be the ultimate goal of IDT’s x86 strategy. By adding a modest processor core to its existing 2-Mbit SRAM, IDT can increase the value of that chip by an order of magnitude. Instead of simply providing the cache for a PC, the company can now provide the CPU/cache subsystem, greatly improving its profit margins.

### Performance Matches Other Low-End Chips

According to Centaur, today’s 200-MHz C6 delivers about the same performance as a 200-MHz Pentium/MMX on the

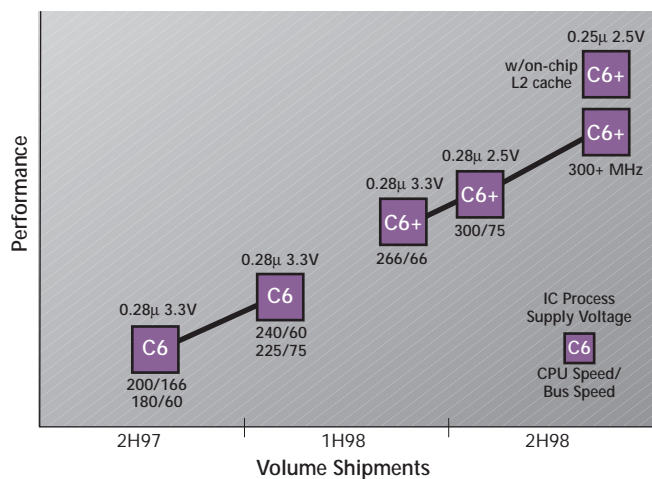



Figure 1. IDT’s C6 roadmap shows faster clock speeds in 1Q98, the improved C6+ in 2Q98, and a 0.25-micron shrink in 2H98 that allows an optional on-chip 256K L2 cache. (Source: IDT/Centaur)

Winstone Business 97 benchmark in a low-cost system configuration. With its integer improvements, the C6+ should outperform the Intel chip by about 7% and match or exceed the performance of AMD’s 200-MHz K6 and Cyrix’s 6x86MX-PR200.

The improvements in MMX and floating-point performance are fairly well matched by those in AMD’s K6 3D and Cyrix’s Cayenne parts. The K6 3D is due in 1H98, about the same time as the C6+, but Cayenne is not expected until 2H98. The new C6+ instructions are aimed at speeding the same 3D applications as the AMD 3D and Cyrix MMXFP instructions, but it is too early to tell if any of the three has a performance advantage over the others.

AMD plans to stay with Socket 7 through the end of 1998 and also expects to add a 256K L2 cache to its K6 3D part in 2H98. This part should be similar to IDT’s integrated offering, although IDT probably can’t match the 400-MHz clock speeds in AMD’s roadmap. Cyrix hasn’t clarified its socket strategy for Cayenne but intimated that it may add a backside bus to Socket 7 or move to a new interface in 2H98.

In summary, the advances in the C6 roadmap should allow it to keep pace with the low-end offerings from AMD and Cyrix. Intel will be moving its Pentium II into the low end of its line by 2H98 (see page 4). This transition will put pressure on all competitors still relying on Pentium/MMX’s Socket 7 at that time.

Given the relatively modest cost of developing the C6, IDT’s goal is to gain only a small (1–2%) share of the x86 market. To do so, it must underprice the competition, namely AMD and Cyrix. IDT would reveal only the single-unit prices for the C6-180 and C6-200, which are \$90 and \$135 respectively. In comparison, AMD sells a K6-200 for \$160 in 1,000-piece quantities. With the C6’s estimated manufacturing cost of \$40, the margins are small but appreciable. For an SRAM vendor, it must seem like a fine way to make a living. 

# PA-8500's 1.5M Cache Aids Performance

## But HP Chip Is Likely to Trail 21264 When It Ships in 2H98

by Linley Gwennap



HP revealed a few more details of the PA-8500 at last month's Microprocessor Forum. The company is still being cagey, however, since the device has not taped out and system shipments are not expected until 2H98. The processor's unique design, which includes a stunning 1.5M of on-chip primary cache, appears likely to fall behind Digital's 21264 in both schedule and SPEC95 performance, but the HP chip may come out ahead when running applications that take advantage of the large cache.

### Fast SRAM With Integrated CPU

As Figure 1 shows, the cache is divided into a 512K instruction cache and two 512K banks of data cache. The caches are four-way set associative, providing a better hit rate than an external direct-mapped cache of the same size. In fact, on some applications, these caches will have hit rates similar to the direct-mapped 2M caches typically used by HP systems today. The data-cache tags are stored in duplicate arrays, allowing snoop transactions to proceed in parallel with cache accesses.

The new chip uses essentially the same CPU core as the current PA-8200, which is built in a 0.5-micron process. Due

to the shrink to 0.25-micron CMOS, that core (excluding the bus interface) now occupies only 26% of the die. At the Forum, PA-8500 designer Bill Queen said his chip will reach at least 360 MHz, about 50% faster than the PA-8200. Due to the magnitude of the process change, we would not be surprised to see the PA-8500 reach 400 MHz or more.

The new core has a few improvements over the PA-8200 (see MPR 10/28/96, p. 18), mainly in the area of branch prediction. The size of the branch history table (BHT) is increased to 2,048 entries, twice the size of the PA-8200's rather meager BHT. The new BHT uses "agrees" mode (see page 22) to improve the prediction accuracy when multiple branches map to the same entry in the BHT. Finally, the size of the TLB is increased by 33% to 160 entries, improving performance on applications, such as transaction processing, that make heavy use of the TLB.

To feed the faster core, HP has doubled the bandwidth of its Runway bus, which it has used in its high-end processors since the PA-7200 (see MPR 3/7/94, p. 12). The new interface transfers data on both edges of a 120-MHz clock, producing a peak bandwidth of nearly 2.0 Gbytes/s. Because the Runway bus is multiplexed, however, the best sustainable bandwidth is just over 1.5 Gbytes/s. Maintaining a 240-MHz data rate will be tricky, but HP has plenty of experience with high-speed system design.

### Integrated Cache Provides Cost Savings

The PA-8500 will set a record by including 130 million transistors, more than 95% of them in the large caches. HP did not disclose the chip's die size, but Queen admitted it is "a bit larger" than current PA-8x00 chips, which measure a hefty 345 mm<sup>2</sup>. The large cache arrays, however, incorporate redundant elements, so they are not susceptible to most defects. As a result, the yield will be at least twice that of the PA-8200. Eliminating the off-chip cache buses reduces the pin count from 1,081 to a more manageable 550. As a result, the PA-8500 will cost about \$160 to manufacture, according to the MDR Cost Model, compared with \$260 for the PA-8200.

The cost savings don't stop there. The on-chip caches eliminate the expensive external caches used by the PA-8200, slashing another few hundred dollars from the system cost.

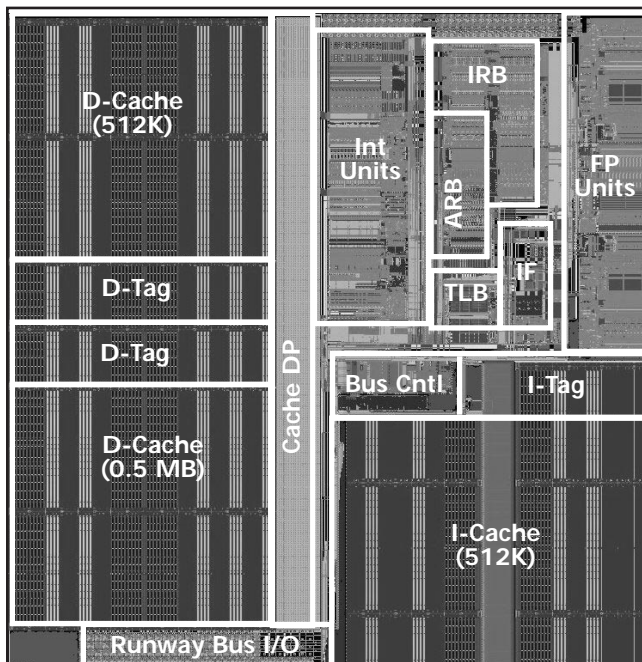


Figure 1. This die plot shows the PA-8500, which will be built in a 0.25-micron four-layer-metal CMOS process using C4 die attach, which eliminates the pad ring. HP did not reveal the die size.

### For More Information

HP does not sell its PA-RISC processors on the open market. For more information on these processors, try [www.hp.com/computing/framed/technology/micropro](http://www.hp.com/computing/framed/technology/micropro).



The PA-8200 uses a complex PC board to route the high-speed cache signals; the new chip is much easier to put on the system board. Thus, the PA-8500 will materially reduce system cost while providing a large performance boost.

Power dissipation will also improve. Queen would not disclose the PA-8500's power consumption but said it is cooler than the current parts, mainly due to a big drop in supply voltage from 3.3 V to 1.8 V. Eliminating the external cache provides further system-level power savings.

HP has not yet revealed the fab for the PA-8500, but since the company has no 0.35-micron capacity, much less the 0.25-micron process needed for the new chip, it is likely to use an outside source. Given its partnership with Intel, that company is a logical choice, but HP might also turn to AMD or to a foundry with leading-edge capacity.

### Not Quite the Last PA-RISC Chip

Over the past summer, HP began dropping hints that it would deploy another PA-RISC processor after the PA-8500, and HP's Queen revealed that chip will be called the PA-8700. Although he declined to provide any details about the device,



MICHAEL MUSTACCHI

**HP project manager Bill Queen explains the benefits of the large primary caches on the PA-8500.**

we suspect it is simply a 0.18-micron shrink of the PA-8500, since most of HP's designers will soon be working on IA-64 chips. The PA-8700 probably won't appear before 2H99, so it is likely to ship after Merced and offer lower performance. Thus, the RISC chip offers HP an insurance policy in case Merced is late or slow, and it offers a crutch to PA-RISC customers who don't want to transition to IA-64 immediately.

In fact, the aging PA-8x00 core appears likely to fall behind Alpha in the performance race as early as next year, despite the PA-8200's current position as the industry's fastest processor. HP says the PA-8500 will deliver 30 SPECint95 and 50 SPECfp95 (base), but given the performance of the

236-MHz PA-8200, clock speeds of 400 MHz or more may be needed to reach these figures.

Given Digital's claim that the 21264 will deliver more than 40 SPECint95 and 60 SPECfp95 (base), HP has backed away from earlier statements that the PA-8500 will deliver industry-leading performance. Queen expects the PA-8500, with its large primary caches, will deliver better performance than the Alpha chip on at least some applications. To prove its point, however, HP must deliver its processor.  $\square$

## Embedded News

*Continued from page 8*

comes with a passive-color LCD controller. The latter part also has more timers and serial channels.

Both parts come in a 176-lead TQFP package (although they are not pin-compatible) and are currently in volume production. In 10,000-unit quantities, the '439 is priced at \$10; the '455 costs \$12. —J.T.

### ■ Digital Swings Across Two New PCI Bridges

Digital Semiconductor (R.I.P.) added two new PCI bridge chips to its already swollen portfolio of such devices. The 21553 and 21554 are the company's first "embedded" PCI bridge chips, fulfilling a particular role within that product category. The two chips are nearly identical; the '553 has a 32-bit PCI interface, while the '554 has a 64-bit bus.

More interesting is what makes these two chips different from normal PCI bridges: instead of a transparent passage between upstream and downstream PCI buses, the '553 and '554 keep the buses separate. In an I<sub>2</sub>O system, the host processor is unaware of devices on the downstream bus and does not map them into its address space during device enumeration. For such systems, the '553 and '554 include an I<sub>2</sub>O messaging unit for communication with the host processor.

In effect, the '553 and '554 allow designers to create a StrongArm-based I<sub>2</sub>O controller similar to Intel's i960RP or

'RD devices (see MPR 6/19/95, p. 10) but without the Intel processor. Or, given recent events (see cover story), perhaps with just a different kind of Intel processor. —J.T.

### ■ Toshiba Spins 74-MHz Windows CE Processor

Toshiba has released one of the first commercial products to come out of its R3900 "Southern Cross" development (see MPR 2/16/95, p. 20), the R3912, a low-power processor for handheld devices. The new chip runs at 74 MHz and dissipates an average of 300 mW, placing it among the more power-efficient 32-bit processors available.

Like virtually all low-power MIPS chips these days, the R3912 includes a hardware MAC unit for soft-modem emulation. The part also includes the TLB required by Windows CE, a 4K instruction cache, a 1K data cache, an IrDA port, and a PCMCIA controller. Toshiba rates the 3.3-V part at 78 Dhrystone MIPS.

The R3912 is a dead ringer for NEC's R4102 (see MPR 4/21/97, p. 4). Both have the same instruction set, cache sizes, Windows CE support, target markets, and \$25 price in 10,000-unit quantities. The Toshiba device is a bit faster, while the NEC chip includes A/D functions. Power dissipation is comparable, given their differences in clock speed. NEC has already scored (its own) handheld PC design win; the Toshiba chip lurks in some Japanese consumer items. Both chips show that MIPS, and Windows CE, are making deeper inroads into the growing consumer market. —J.T.  $\square$

# Gshare, “Agrees” Aid Branch Prediction

## *New Algorithms Have Minimal Hardware Cost But Improve Accuracy*

by Linley Gwennap

Several recently announced microprocessors have adopted one (or both) of two new algorithms for branch prediction. The first, known as Gshare, is a method of indexing into a large branch history table (BHT). The second, “agrees” mode, is a method of encoding the history bits themselves. Although both methods offer modest improvements in branch-prediction accuracy, neither has significant hardware cost compared with current methods. With modern microprocessors facing lengthier branch penalties, both techniques are likely to become standard practice as designers revise their processors over the next couple of years.

### Gshare Improves Two-Level Algorithm

The Gshare algorithm is a variation of the two-level algorithms first published by Yeh and Patt (see MPR 3/27/95, p. 17). The complete two-level algorithm uses a table of branch patterns to index into a table of history values. Because this algorithm requires a time-consuming pair of lookups, commercial processors have generally used a simplified version in which a global history value is used to index into the history table.

To allow some amount of per-branch history, this global history is typically concatenated with some bits of the branch address. A paper by Digital’s Scott McFarling ([www.research.digital.com/wrl/techreports/abstracts/TN-36.html](http://www.research.digital.com/wrl/techreports/abstracts/TN-36.html)) calls this method Gselect. It improves performance over an index that ignores global branch history due to the correlations between nearby branches. For example, in the sequence

```
IF (x<1)...
IF (x>1)...
```

the direction of the second branch is clearly influenced by the outcome of the first branch.

For a BHT with a fixed number of entries  $e$ , the size of the index is  $\log_2(e)$  bits. For example, a 1,024-entry BHT requires a 10-bit index. Using Gselect, these bits must be divided between the global history and the branch address, for example, 5 bits of each.

McFarling notes that the indexes generated by Gselect have a lot of redundancy, since few combinations are relevant. Thus, hashing the global history and the branch address using an XOR operation will typically not remove useful information. Prediction is improved because more bits of the global history and the branch address can be used. In the 1,024-entry BHT, up to 10 bits of each could be combined to create the index. McFarling calls this method Gshare.

Extending the global branch history beyond several bits does not significantly improve accuracy and can instead

introduce noise into the hashed index. Thus, for a large BHT, the index may use only 4–8 bits of global history XORed with the upper bits of the branch address.

The hardware cost of implementing Gshare instead of Gselect is negligible: possibly extending the size of the global branch history register by a few bits and adding a set of XOR gates to the index logic. According to McFarling, Gshare improves accuracy by less than 1% over Gselect, but the cost is so small that there is no reason not to use the better method.

### Agrees Mode Avoids Contention

A typical BHT uses two bits to encode the history of a particular branch (see MPR 3/27/95, p. 17). These bits indicate if the branch has been taken or not taken. The PA-8500 (see page 20) uses a different encoding: the bits indicate whether the outcome of the branch has agreed or disagreed with the static prediction. In PA-RISC, branches include a static prediction bit in the instruction itself.

At first glance, there appears to be no practical difference between the two methods. Because most BHTs have no tags, however, multiple branches might map to the same BHT entry. Sometimes, a new branch simply takes over an entry from an old branch. In the worst case, two active branches can both be updating a single BHT entry.

In the taken/not taken scheme, there is a 50% chance that the two conflicting branches will go in opposite directions, causing frequent mispredictions. If both branches are following their static predictor, however, the agrees method will correctly predict both branches even though they map to the same BHT entry. The chances of a misprediction are reduced to  $2p(1-p)$ , where  $p$  is the success rate of the static predictor. For  $p=0.7$ , the agrees mode mispredicts only 42% of the time in situations where two branches map to the same BHT entry. While this improvement is small, the hardware cost of implementing agrees mode is essentially zero: since only the encoding is changed, no new storage is needed.

Agrees mode is an obvious solution for instruction sets—like PA-RISC, PowerPC, and MIPS—that include static branch prediction. Centaur found a way to make it work with x86, which does not. The C6+ (see page 17) generates a static prediction based on the opcode and direction (forward or backward) of the branch, then compares this prediction with the actual result to encode the branch history using agrees mode. This on-the-fly prediction adds hardware cost, but static prediction can be done very simply (the simplest way is backward taken, forward not taken), allowing agrees mode to be applied to any instruction set. Thus, like Gshare, agrees mode is likely to become widely used.  $\square$

# IBM's Power3 to Replace P2SC

## Power3 to Complete IBM's Transition to PowerPC Architecture

by Peter Song



To complete its transition from POWER to PowerPC, IBM has announced the Power3 processor, a 64-bit PowerPC chip that will replace the P2SC (see MPR 8/26/96, p. 14) in its RS/6000 workstations and servers in 2H98. At last month's Microprocessor Forum, Mark Papermaster of IBM Microelectronics described IBM's latest design, which has its heritage in both the ill-fated PowerPC 620 and P2SC. As the P2SC runs out of gas, reaching only 160 MHz in a 0.25-micron process, IBM plans to replenish its RS/6000 product lines with Power3 chips, which will eventually reach 500 MHz in more advanced processes.

Executing two multiply-add operations in each cycle, Power3 maintains the emphasis on floating-point performance that has become the trademark of high-end RS/6000 systems. Executing two integer and two load-store operations—twice the rate of the P2SC—in each cycle, it also improves that chip's poor SPECint95 performance by 35% at

equivalent clock speeds. Although its 64K data cache is only half the size of the P2SC's, its advanced core, a dedicated L2 cache, and aggressive prefetching mechanisms more than make up for the smaller cache.

Power3 is the processor that will complete IBM's transition from the POWER to the PowerPC architecture. Among its many advantages, the PowerPC architecture offers better multiprocessor scalability, while enabling simpler designs for high-end systems than possible with the POWER architecture. It also brings a 64-bit instruction set to a growing segment of multiprocessor servers that demand huge amounts of memory. Its 64-bit format will become even more useful when multimedia instruction-set extensions, which the PowerPC alliance has been quietly developing, are added to PowerPC processors.

### Modified PowerPC 620 Core

The Power3 design team started with the largely finished 620 design (see MPR 10/24/94, p. 12) for the 64-bit PowerPC core, backside cache interface, and PowerPC 6xx bus. The team added the second floating-point and load-store units, as Figure 1 shows, making Power3 better suited for applications that combine floating-point math with large data sets. The team also added more queues and rename registers, increasing to 32 the number of instructions that can be in process. The two additional execution units give Power3 a peak execution rate of eight instructions per cycle. The team reduced the sustainable execution rate from the P2SC's six instructions per cycle to four per cycle, limited by the decode/dispatch unit, but the reduction is unlikely to affect overall performance.

The team made significantly more changes to the 620 chip's memory subsystem to sustain two loads and two floating-point operations per cycle. The team doubled the data-cache size to 64K, improving the cache-hit rate. The cache is

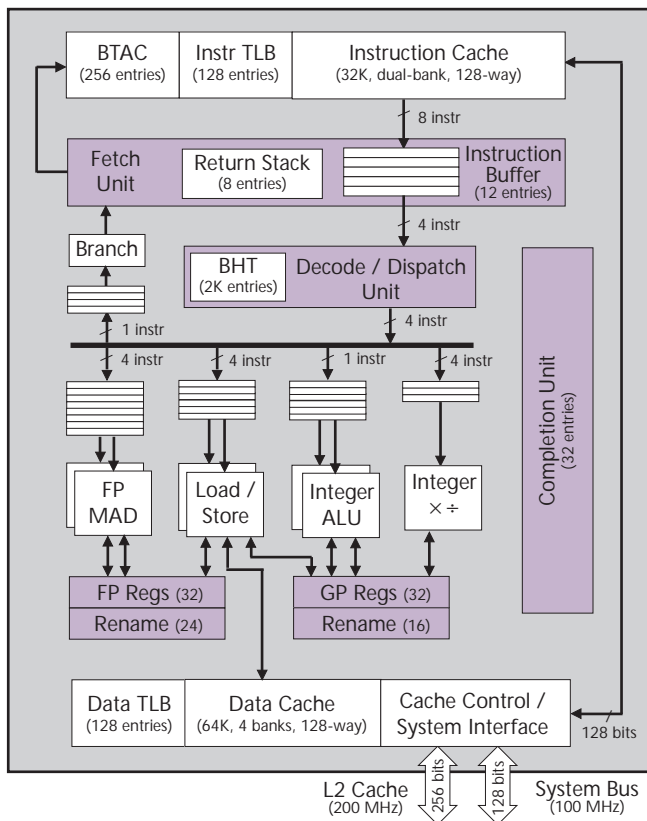


Figure 1. Power3 can execute two load and two integer or floating-point instructions in each cycle.

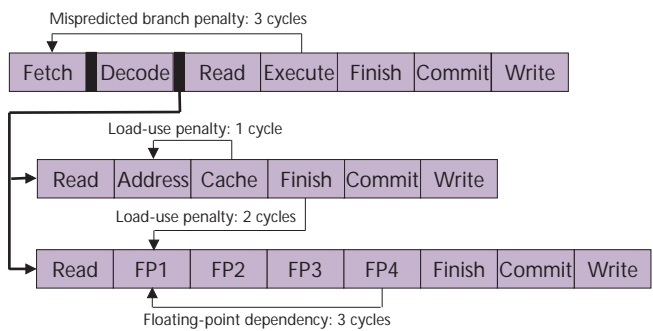


Figure 2. Power3's short pipeline keeps its branch-misprediction penalty to three cycles. Instructions incur possible delays, shown with black bars, in the instruction buffer and the issue queues.



interleaved into eight banks on double-word boundaries, reducing conflicts among the two loads and a store that can access the cache in each cycle. Taking advantage of the greater localities found in large engineering and scientific data sets, the team doubled the line size to 128 bytes, increasing the L2 interface to 32 bytes wide to boost bandwidth and accommodate the larger line size.

### Power3 Uses a Short Pipeline

Unlike competitive chips such as Digital's 21264 (see MPR 10/28/96, p. 11) or HP's PA-8500 (see page 20), which use five pipeline stages before instructions enter the first execution stage, Power3 keeps this front end of the pipeline short, using only three stages. Taking advantage of its lower clock speed, Power3 needs only one cycle to access the instruction cache, one cycle to decode and dispatch the instructions to different execution units, and one more cycle to access the operands, as Figure 2 shows. Power3's relatively short pipeline keeps its mispredicted branch penalty to only three cycles, 2–4 cycles shorter than its competitors'.

Up to eight instructions are fetched from any position within a 128-byte cache line and placed into the 16-entry instruction buffer. The large line size and the word-aligned cache access increase the frequency of fetching eight useful instructions in each cycle. The instruction cache consists of two banks of fully-associative arrays, yielding 128 sets in each array. The dual-bank design supports both instruction fetch and cache reload to different banks in the same cycle. All 128 bytes of a cache line are updated in a single cycle, further reducing the number of cycles the cache is unavailable for instruction fetch.

In each cycle, up to four instructions in the instruction buffer are decoded and dispatched in program order to the execution units. Except for the branch and multicycle-integer units, which can accept at most one instruction per cycle, each execution unit can accept up to four instructions per cycle, provided that the unit has enough available slots in its issue queue. The instructions within the buffer are shifted by zero, two, or four slots, ensuring that at least three instructions are available for dispatch in each cycle. Supporting all possible shift options would have doubled the number of inputs to each entry.

Up to eight instructions—two floating-point, two load/store, two single-cycle integer, a multicycle integer, and a branch—can begin execution in each cycle. Ready instructions are issued out of order from the issue queues, allowing instructions of different types, as well as of the same type, to execute out of order. The load/store and branch instructions are issued in program order, simplifying the design. Unlike the 620's reservation stations, Power3's issue queues do not

store the operands, requiring an additional cycle to access the operands in most cases. When the queues and the execution pipelines are empty, such as when the machine is recovering from a mispredicted branch or processing an exception, instructions read the operands during the dispatch cycle, bypassing the issue stage.

When an instruction finishes execution, its result is written to a rename register during the finish stage. The instruction's execution status—that it finished execution and did or did not detect an exception—is written to the completion buffer. Up to four instructions can be retired in the commit stage. The results from retired instructions are copied from the rename registers to the architectural registers in the write stage. Because execution results are made available in the rename buffers as soon as they are produced, the last three pipeline stages have little effect on performance as long as enough rename registers are available.



**Mark Papermaster describes Power3's short pipeline that keeps branch penalty low.**

MICHAEL MUSTACCHI

### Only a Few Branches Are Predicted

The PowerPC branch model allows Power3 to execute branch instructions as early as the decode stage, reducing the number of branch instructions that are predicted. In contrast to the *compare-and-branch* model used in the PA-RISC and MIPS architectures, the POWER and PowerPC architectures support the *condition-code-and-branch* model. This model separates the instructions that generate the branch condition from the instructions that change the program flow, allowing the branch condition to be generated in advance. In addition, the POWER and PowerPC architectures provide eight sets of condition codes, allowing multiple branch conditions to be precomputed.

For a taken branch, the branch target address is also needed to redirect the fetch stream. In the decode stage, the program-counter-relative branch instructions readily provide the target address, but the register-direct branch instructions do not. Instead of a general-purpose register, these branch instructions in the POWER and PowerPC architectures use one of two special registers (LINK and COUNT registers). Because these special registers are much smaller than the register file and the instructions that update them are easily detected, they are faster and easier to access than a general-purpose register, especially in superscalar designs.

Many POWER and PowerPC designs take advantage of these features to execute branch instructions as soon as the branch conditions are known. In Power3, each group of 4-bit condition codes is mapped to a rename register, allowing speculative and out-of-order generation of multiple branch conditions. Most instructions that modify the LINK or COUNT registers, such as BRANCH-AND-LINK, DECREMENT-AND-BRANCH, or MOVE-TO-LINK, can also execute speculatively. These special

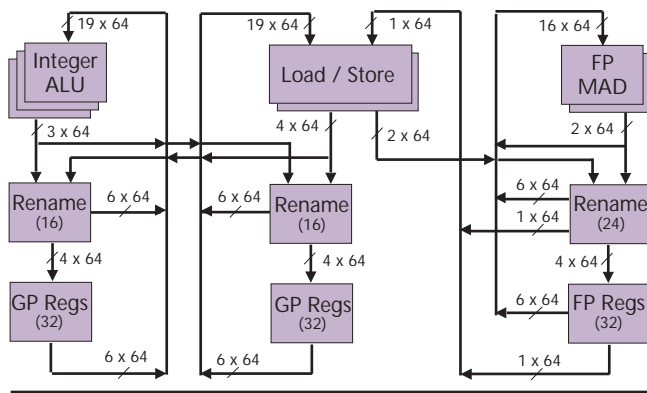


Figure 3. Power3's core requires many 64-bit buses to read operands and return execution results. The number of operand-read ports is reduced from 12 to 6 by using two sets of registers.

registers, along with their rename registers, are made accessible for executing a branch instruction in the decode stage.

### BTAC Predicts Branch Condition, Target Address

For branch instructions whose conditions are not known in the decode stage, Power3 uses a 2,048-entry BHT (branch history table) to predict their branch direction. Because a branch is often resolved in the decode stage or soon thereafter, the benefit of the BHT when used to predict the current encounter of the branch is less in Power3 than in designs with deeper pipelines. To better utilize the BHT, however, Power3 uses the BHT to predict both the current and the next encounter of each conditional branch, using a BTAC (branch target address cache).

A BTAC is generally used to predict the target address of a predicted-taken or, in some designs, a known-taken branch instruction. It avoids a register-file access and an address calculation in many designs. Its prediction is highly accurate because branch instructions tend to jump repeatedly to the same target instruction, even in object-oriented programs that use dynamic binding.

In Power3, the 256-entry BTAC is used both to predict the target address of a branch and to predict the branch direction (taken or not-taken). By keeping only predicted-taken branch instructions in the BTAC, a branch that misses the BTAC is naturally predicted not taken. As each conditional branch is resolved in the execute stage, the instruction is added to the BTAC if its two-bit branch history changes from not taken to taken, effectively predicting the branch direction for its next encounter. Conversely, the instruction is deleted from the BTAC if its history changes from taken to not taken.

The BTAC is accessed in the fetch stage, resulting in a zero-cycle delay for taken-branch instructions when correctly predicted. It is accessed with the fetch address, not with the specific address of a branch, since the address of the branch instruction is unknown. In effect, the BTAC is accessed in case one of the eight instructions being fetched is a predicted-taken branch. When the access hits the BTAC—indicating

	Power3 (200 MHz)	PA-8500 (360 MHz)	21264 (600 MHz)
64-bit Integer Mul	9 cycles	n/a	7 cycles
64-bit Integer Divide	37 cycles	n/a	>64 cycles
FP Mul-Add (DP)	4 cycle	3 cycles	4 + 4 cycles*
FP Add/Mul (DP)	4 cycle	3 cycles	4 cycles
FP Divide (SP)	14 cycles	18 cycles	12 cycles
FP Square Root (SP)	14 cycles	18 cycles	16 cycles
FP Divide (DP)	18 cycles	32 cycles	16 cycles
FP Square Root (DP)	22 cycles	32 cycles	33 cycles

Table 1. Power3's execution latencies are comparable to those of its competitors but at a lower clock speed. n/a = not applicable. \*using multiply and add instructions. (Source: vendors)

that one of the instructions being fetched is a predicted-taken branch—the returned target address is used for accessing the instruction cache and the BTAC in the next cycle.

Other processors provide a zero-cycle delay for taken-branch instructions when correctly predicted. Instead of using a BTAC, the 21264 keeps a 12-bit field for each 32-byte line, providing the next-line and next-set prediction for each instruction-cache access. This design uses fewer bits per branch instruction than does Power3, since it does not need a separate tag array and uses only a 12-bit index instead of the virtual addresses for the branch and the target instructions.

### Power3 Sustains Four Instructions Per Cycle

Power3 uses rename registers for the GPRs (general-purpose registers), FPRs (floating-point registers), and the CCR (condition-code register) to allow out-of-order and speculative execution of most instructions. The few exceptions are stores and certain move-to-special-register instructions that are difficult to undo. As instructions are dispatched to the issue queues in program order, their destination registers are mapped to the rename registers, allowing out-of-order access to the registers. Although instructions can be issued out of order from the issue queues and thus their operands can be read out of order from the register files, the rename registers eliminate anti- and output-dependency hazards by enabling the register files to be updated in program order.

As Figure 3 shows, Power3 uses two copies of the GPRs and the general-purpose rename registers, reducing the number of read ports needed to support the five execution units from 12 to 6. One copy provides two operands to each of the three integer units, and the second copy provides three operands—two for the address operands and one for the store data—to each of the two load-store units. The 21264 also uses two copies of integer registers but assigns a copy to a pair of load-store and integer execution units. Due to its short cycle time, the 21264 requires an extra cycle to transmit a result from one pair to the other. Operating at a lower speed, Power3 does not, simplifying the dispatch logic.

The rename registers have four more read ports for updating the GPRs with the results of retired instructions, matching the sustainable execution rate of four instructions per cycle. The rename registers also have seven write ports:

one from each of the integer units and two from each of the load-store units. The second write port for the load-store units allows them to execute a LOAD-WITH-UPDATE instruction, which returns both load data and an updated address value in each cycle.

The 21264 can execute two floating-point instructions per cycle, but only when one is a multiply and the other is an add. In contrast, Power3 has two identical FPUs, delivering twice as many floating-point results per cycle as the 21264. Power3's FPUs execute multiply-add instructions, as Table 1 shows, taking only half the number of cycles as the 21264 to calculate the common  $A \times B + C$  operation. At only a third of the Alpha chip's 600-MHz clock speed, however, Power3 is unlikely to deliver better overall floating-point performance. HP's PA-8x00 can also execute two multiply-accumulate instructions in each cycle but takes only three cycles to produce the results.

### Hardware Prefetch Reduces Cache Misses

Power3 uses instruction- and data-prefetch mechanisms to reduce pipeline stalls due to cache misses. The instruction cache is two-way interleaved on cache-line boundaries, allowing one bank to be accessed for instruction fetches while the other bank is accessed for the next cache line. When the former access hits in the cache but the latter access does not, a prefetch request for this next cache line is issued to the L2 cache. Because the prefetch is still speculative, the request is not propagated to the main memory if it misses in the L2 cache, allowing the request to be canceled upon detecting a mispredicted branch instruction. An instruction prefetch takes six cycles from the 200-MHz L2 cache.

For the data cache, a prefetch mechanism keeps track of up to four prefetch streams and, for each stream, fetches up to two cache lines ahead from the L2 cache or main memory. It identifies a prefetch stream only if the accesses in the stream advance from one 128-byte block of memory to the adjacent—sequentially succeeding or preceding—block of memory. The prefetch mechanism is therefore restricted to a stride of one cache line.

To establish a prefetch stream, the prefetch mechanism monitors every access that misses in the data cache, searching for cache-miss references to two adjacent 128-byte blocks of memory. Upon finding such a pair, it initiates a prefetch request for the next 128-byte block. The address of the prefetch request, along with the ascending or descending prefetch direction, is kept in a four-entry prefetch queue. Once a prefetch stream is identified, the address of every data-cache access is checked with the addresses in the prefetch queue. When a match is found, a prefetch request for the next 128-byte block is made, and the address in the matching entry is updated with the address of the new prefetch request.

Using this prefetch mechanism, Power3 executes the inner loop of DAXPY benchmark at the rate of 27 cycles per iteration, according to IBM. This loop contains 32 load and

16 store operations, which require 26 cycles for Power3 to execute, assuming no data-cache misses. With the prefetch mechanism disabled, Power3 executes the same loop in 36 cycles.

The Power3's prefetch mechanism is likely to be less effective, however, in programs that are not as well behaved as DAXPY. Requiring two cache-miss references to identify a prefetch stream and limiting the prefetch stride to 128 bytes, the mechanism's effectiveness depends largely on carefully organizing the data sets. UltraSparc-3's prefetch cache (see MPR 10/27/97, p. 29), in contrast, places no restriction on the prefetch stride, allowing complete freedom in organizing the data sets. In addition, the prefetch cache relies on software to identify the load instructions whose data should be prefetched, simplifying the prefetch hardware while incurring fewer cache-miss references than Power3.

### Chip's Interfaces Cap CPU Speeds at 200 MHz

Power3 is built initially in IBM's CMOS-6S2 process, a hybrid of 0.25-micron transistors with 0.35-micron metal layers (see MPR 9/16/96, p. 11), and occupies a 270-mm<sup>2</sup> die. It uses the same 1,088-pin ceramic package as the P2SC but a different pinout. The MDR Cost Model estimates the manufacturing cost of the Power3 chips to be \$160, only 40–60% of the competitors'. For these nonmerchant chips, however, the cost advantage provides no real benefit.

According to IBM's Papermaster, current Power3 chips are functional, running at an excess of 200 MHz, and will be available in systems in 2H98. Operating the cache and system interfaces at the full speed and half the speed of the processor, respectively, the 200-MHz Power3 chips deliver 6.4 Gbytes/s of cache bandwidth and 1.6 Gbytes/s of memory bandwidth. At 200 MHz, IBM estimates Power3 will deliver 28 SPECfp95 and 12 SPECint95 (base). By the time Power3 ships, it will compete against the 21264 and PA-8500, which are slated to deliver 30–40 SPECint95 and 50–60 SPECfp95 (base).

Although current Power3 chips may operate at speeds faster than 200 MHz, Power3 systems are unlikely to be offered at the faster speeds, because the current design does not support fractional processor-to-cache and processor-to-system clock ratios (i.e., 3:2 mode). Faster Power3 chips require either using correspondingly faster SRAMs in 1:1 mode or operating the cache interface in 2:1 mode, half the speed of the processor. The former solution is expensive and feasible at speeds only up to 250 MHz, while the latter delivers lower performance at speeds below 400 MHz. At speeds between 200 and 250 MHz, the system interface presents a dilemma that conflicts with the cache interface; it can no longer operate in 2:1 mode but degrades performance in 3:1 mode.

IBM is already working to remedy the situation, first by migrating the Power3 design to its CMOS-7S process (see MPR 8/4/97, p. 14). In this 0.2-micron process, which also uses copper interconnects, IBM expects first silicon of 350-MHz Power3 processors in 2Q98. The die size will shrink to 160 mm<sup>2</sup>, with a few additional functions. Although the process will be in production by 2H98, the belated design will



## Prices & Availability

Power3 will not be sold as a standalone product. Systems using Power3 will ship in 2H98. For more information, contact your local IBM sales office or the IBM RS/6000 Web site [www.austin.ibm.com/hardware](http://www.austin.ibm.com/hardware).

keep faster Power3 chips unavailable in systems until 1H99. In 2H99, IBM plans a second derivative of Power3 in a true 0.18-micron process, hoping to reach speeds up to 500 MHz.

The faster Power3 chips will support fractional bus modes—5:2 and 7:2 for processor-to-bus and 3:2 for processor-to-cache interfaces—which will allow the core to run at its full speed. Using a set-prediction mechanism, the new chips will also support a four-way set-associative L2 cache without using additional pins for the data bus, similar to the two-way set prediction used in SGI's R10000 (see MPR 10/24/94, p. 18). To select the correct set in a single cycle when the prediction is wrong, however, the chips will have 100 additional pins to read all four tags simultaneously.

We believe bringing the L2 tag array on chip would offer performance advantages over adding more pins to support a better L2 cache. In addition to better supporting a set-associative L2 cache, the on-chip tag array would also reduce the L2-miss and snoop-response latency cycles. Because the array could be clocked at the CPU speed, not the L2-cache speed, it also offers more bandwidth for the CPU and the coherency traffic, making the design more suitable in large-scale MP configurations. The tag array for a 16M L2 cache would be comparable in size to the current Power3's 64K data cache, which occupies about 27 mm<sup>2</sup> (10% of the die size) in the current design. Since SRAMs are 2.3× denser in CMOS-7S than in CMOS-6S2, the array would occupy only about 11 mm<sup>2</sup> in the faster Power3, a reasonable area overhead for gaining the extra performance.

### Power3 Completes IBM's Transition to PowerPC

Due to its "brainiac" design (see MPR 3/8/93, p. 3), Power3 operates at much lower speeds than its competitors, such as the PA-8500 and 21264 processors. These competitors are likely to operate at 1.5–3× the speeds of Power3, using a comparable process. Even with 30–50% fewer logic transistors, they can also process more instructions out of order than does Power3.

As Figure 4 shows, the Power3's caches occupy only 15% of the die area. In comparison, the 1.5M caches in the PA-8500 occupy nearly 60% of the die area. Within the Power3's heritage, the percentage of total transistors used for logic has steadily increased—from 20% in Power2 to 38% in the P2SC to now 47% in Power3. With wire delays taking an increasingly large portion of cycle times as process dimensions shrink, using more of the available transistors in logic is likely to make the wire-delay problem worse. Considering

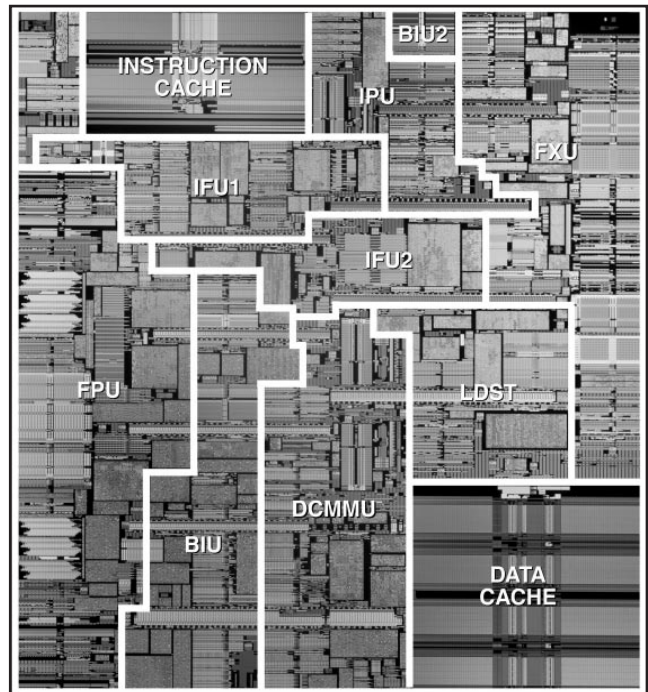



Figure 4. The Power3's two caches occupy only 15% of the 270-mm<sup>2</sup> die, which is built in a hybrid 0.25-micron process. In comparison, the 1.5M caches in the PA-8500 occupy 60% of its die area.

that migrating the P2SC from the 0.29-micron CMOS-6S to CMOS-6S2 increases its speed by only 20%, IBM is likely to face tough challenges in delivering Power3 derivatives that meet their clock-speed goals.

Higher clock speed is not the same as higher performance, however, especially for many server applications that have low hit rates for the on-chip caches. For these cache-busting applications, hiding memory latency and providing sufficient memory bandwidth is as important as high clock speed, a point the Power3 designers did not miss. For two processors that deliver comparable performance at different clock speeds, the processor with a lower clock speed should tolerate cache misses better, due to fewer memory-latency cycles. The ability to reorder a large number of instructions helps, but that ability cannot completely hide memory latencies that range up to more than a hundred cycles.

When IBM replaces the P2SC processor with Power3 in the RS/6000 product lines, it will complete the long-awaited transition from the POWER to the PowerPC instruction set, establishing a single architecture for IBM's entire computer product lines (except legacy mainframes and PCs). The transition is already completed in the AS/400 minicomputers, which use the 64-bit PowerPC A10 and A30 processors (see MPR 7/31/95, p. 15). Although many RS/6000 workstations and servers are already powered by PowerPC processors, the high-end servers still use the P2SC for its superior floating-point performance. With Power3 in high-end server systems, IBM will fulfill its vision of providing "palmtops to teraflops" using a single architecture. 

# RISC on the Desktop: Game Over

## *Sun the Only Holdout in IA-64 Sweep*



With Digital's commitment to build a full range of systems based on IA-64 processors, Intel's new architecture—though still a paper tiger—has nearly completed its sweep of the computer industry. Unless Intel blunders in some major way, it seems inevitable that its microprocessor

dominance will gradually be extended to include workstations and servers as well as PCs.

The list of companies signed up to build IA-64 systems is impressive. HP, of course, has been on board from the beginning. Silicon Graphics revealed two months ago that it would build Intel-based systems, though it has refrained from publicly committing to IA-64 specifically. The three largest PC makers—Compaq, Dell, and IBM, which stated its commitment to the architecture at last month's Microprocessor Forum—are all aboard the IA-64 train. So are Bull, NCR, Sequent, Stratus, Hitachi, NEC, Unisys, and ICL.

Many of these companies have existing Intel-based product lines, and some of them have RISC lines as well. One could argue that these companies are just continuing the evolution of their Intel lines, and that this does not necessarily affect the RISC-based products. But an x86 product line and a RISC product line have clearly distinct performance positions. With IA-64 systems, the performance gap will be much smaller and, in time, the IA-64 systems probably will pull into the lead.

A very high-end focus is the only apparent survival strategy for the RISCs: build something that doesn't overlap with anything Intel builds. But in time, the space above the fastest IA-64 processors may become very small.

RISC-based systems won't disappear instantly. The RISC systems have unique software and customer bases that will take time to migrate. In addition, it probably won't be until the second-generation IA-64 processor ships in 2001 that the architecture will really shine. It will take time for the compilers and operating systems to mature, as well as for applications to get ported. But eventually, unless there turns out to be a fundamental flaw in the IA-64 approach, the architecture is likely to develop the same kind of momentum that drove the x86 architecture to dominance.

Despite many years of campaigning, none of the RISCs has large customers beyond the architecture owner (except for embedded applications). None has a customer base that approaches the weight of the companies signed up for IA-64.

Digital executives claim that the company's plans for Alpha remain unchanged, and that they still hope to drive it

into the mainstream PC market (see cover story)—but this just doesn't seem realistic. Alpha has made little headway against x86, and it will have a lot more difficulty making headway against x86 and IA-64. The publicity surrounding the Intel/Digital deal has been damaging, and Digital's Alpha customers must be very skeptical. Digital's system business will find itself increasingly torn between Alpha and IA-64.

Silicon Graphics plans to stick with MIPS for its high-end systems but has no expectations of driving it into lower-cost desktops. The current turmoil at SGI could lead to an acceleration of its move to Intel-based systems.

Sun is the last IA-64 holdout among major computer companies. Sun's strategy is anti-Intel, anti-Microsoft: SPARC, Unix, and Java. When I asked whether Sun would build IA-64 systems, Scott McNealy unambiguously replied, "absolutely not." But Sun has committed to porting Solaris to IA-64, making it easy to switch when the time is right. I'd expect McNealy to deny that Sun has any plans to build IA-64 systems right up to the day Sun announces them.

In the old days of the computer business, companies were vertically integrated and competed in all levels of technology. The PC business established a different model, in which a few technology suppliers feed hundreds of computer companies, which differentiate through system design, peripherals, packaging, marketing, distribution, and support. While there won't be one product, as with the IBM PC, that establishes IA-64's role, the number of companies supporting it—and the resources Intel can put behind it—are likely to create a critical mass that no RISC architecture has been able to achieve.

IA-64 clearly has the potential to have a chilling effect on competition in high-end microprocessors. It may or may not be a real breakthrough—but it doesn't have to be. It just has to be pretty good. The danger that progress will slow, and that Intel won't price as aggressively, in the absence of vibrant competition is real—but the price of the free-market system is that strong companies may get stronger. The industry must depend on the Department of Justice to make sure that Intel plays by the rules.

Innovation in computers will continue, but its focus will shift from instruction-set architecture and microprocessor design to delivering system-level solutions. Companies that want to compete in the microprocessor industry have two realistic approaches: focus on embedded applications, or build processors that are software-compatible with Intel's. ■

See [www.MDRonline.com/slater/IA64](http://www.MDRonline.com/slater/IA64) for more on this subject. I welcome your feedback at [m Slater@mdr.zd.com](mailto:m Slater@mdr.zd.com).

## RECENT IC ANNOUNCEMENTS

PART NUMBER	VENDOR	DESCRIPTION	PRICE/QUANTITY	AVAILABILITY
<b>MICROPROCESSOR</b>				
VR4111	NEC 800.366.9782	MIPS processor runs Windows CE at 80, 100 MHz; uses MIPS-16 code compression, with 32-bit external bus, A/D, IrDA, PCMCIA control.	\$25/10,000	Samples—Now Prod.—1Q98
ST6230	SGS-Thomson 617.259.0300	Microcontroller has 8-bit ST62 core, 20 I/O pins, 8-bit timer, 16-bit auto-reloading timer, serial-peripheral interface (SPI), 8-bit A/D converter, more.	\$1.95/100,000	Prod.—Now
ST6232	SGS-Thomson 617.259.0300	Microcontroller has 8-bit ST62 core, 30 I/O pins, 8-bit timer, 16-bit auto-reloading timer, serial-peripheral interface (SPI), 8-bit A/D converter, more.	\$5.13/1,000	Prod.—Now
ST6235	SGS-Thomson 617.259.0300	Microcontroller has 8-bit ST62 core, 36 I/O pins, 8-bit timer, 16-bit auto-reloading timer, serial-peripheral interface (SPI), 8-bit A/D converter, more.	\$4.43/1,000	Prod.—Now
KS86x6104	Samsung 408.954.7000	Microcontroller for USB mouse or joystick has 8-bit SAM87 core, 4K of OTP or mask ROM, 144 bytes RAM, 41 instructions, and 16 registers.	\$1.50/1,000	Samples—Now Prod.—1Q98
KS86x6008	Samsung 408.954.7000	Microcontroller for USB keyboards has 8-bit SAM87 core, 4K/8K of mask/OTP ROM, 224 bytes RAM, 41 instructions, and 16 registers.	\$2.00/1,000	Samples—Now Prod.—1Q98
KS88x6016/24/32	Samsung 408.954.7000	Microcontroller for USB hubs has 8-bit SAM87 core, 12-Mbps isochronous interface, 8-bit, 8-channel PWM output, 8-bit A/D converter, DDC, more.	\$6.00/1,000	Samples—Now Prod.—1Q98
<b>INTERFACE</b>				
CS5360	Crystal 512.912.3113	Low-cost 24-bit audio A/D converter has peak-level-detect feature; pin-compatible with 20-bit CS5335 converter; runs from 5-V supply.	\$7.50/1,000	Samples—Now Prod.—1Q98
CS4390	Crystal 512.912.3113	Stereo 24-bit audio delta-sigma D/A converter has 106-dB dynamic range, 115-dB signal/noise ratio, and -97-dB THD+N; compatible with CS4329.	\$5.30/1,000	Prod.—Now
CS4925	Crystal 512.912.3113	Audio decoder handles both Dolby Digital (AC-3) and 5.1-channel MPEG-2; microcoded chip built in 0.35-micron process; in PLCC-44.	\$15/100,000	Samples—Now Prod.—1Q98
PCI1210 PCI1221	Texas Instruments 800.477.8924	PCI-to-CardBus controllers support one (*10) or two (*21) CardBus slots; both are compatible with 3.3-V and 5-V signaling; in 144- and 208-TQFP.	\$9/1,000	Samples—Now Prod.—1Q98
<b>MEMORY</b>				
M34C02	SGS-Thomson 617.259.0300	Small E <sup>2</sup> PROM is specifically designed for serial presence detect on 168- and 200-pin DIMMs; with 2-Kbit capacity, I <sup>2</sup> C interface, software locking.	\$0.50/1,000	Prod.—Now
CY7C42xx	Cypress 408.943.2600	Synchronous FIFOs have 512-Kbit or 1-Mbit capacity, 100-MHz frequency, larger capacity than other synchronous FIFOs.	\$35.50/10,000	Prod.—Now
<b>MISCELLANEOUS</b>				
CY2308-xx	Cypress 408.943.2600	Clock buffers use internal PLL for "zero delay" propagation of 100-MHz signals; optional multiply input reference by 2x or 4x.	\$2.45/100,000	Prod.—Now
SDA9388	Siemens 800.777.4363	Chip provides low-cost picture-in-picture for televisions; incorporates digital color decoder, filters, A/D and D/A, and V-chip functions.	\$7/100,000	Samples—1Q98 Prod.—2Q98
TCM37C13A TCM37C14A	Texas Instruments 800.477.8924	Pulse-code modulated (PCM) codes with voice-band filtering have programmable transmit and receive gain; for PABXs, line cards.	\$2.02/1,000	Samples—Now Prod.—1Q98



by Rich Belgard, Contributing Editor

The following U.S. patents related to microprocessors were issued recently. Please send comments or questions via e-mail to [belgard@umunhum.stanford.edu](mailto:belgard@umunhum.stanford.edu).

#### 5,628,024

*Computer architecture capable of concurrent issuance and execution of general-purpose multiple instructions*

Issued: May 6, 1997

Inventor: Robert W. Horst

Assignee: Tandem

Filed: June 7, 1995

Claims: 6

A system for issuing a family of instructions during a single clock. Logic attached to the instruction decoder determines whether resource conflicts would occur if the family were issued during one clock. If no resource conflicts occur, an execution unit executes the family regardless of whether dependencies among the instructions in the family exist.

#### 5,628,021

*System and method for assigning tags to control instruction processing in a superscalar processor*

Issued: May 6, 1997

Inventors: Kevin R. Iadonato, et al

Assignee: Seiko Epson

Filed: April 4, 1994

Claims: 62

In an out-of-order microprocessor, a register file stores data for each instruction. A queue contains instructions and tags. The tags are arranged in the queue in program order. Data for instructions can be read out of the register file in program order based on the tags.

#### 5,627,983

*Processor architecture providing out-of-order execution*

Issued: May 6, 1997

Inventors: Valeri Popescu, et al

Assignee: Hyundai

Filed: June 6, 1995

Claims: 18

An out-of-order processor that includes a shelving unit for temporarily storing instructions to issue and to temporarily receive results of execution; functional units to perform the instructions; and a retirement unit to commit instructions.

#### 5,627,985

*Speculative and committed resource files in an out-of-order processor*

Issued: May 6, 1997

Inventors: Michael A. Fetterman, et al

Assignee: Intel

Filed: January 4, 1994

Claims: 33

An out-of-order processor that contains, and interconnects, a register alias table, a reorder buffer, and a reservation station.

#### 5,627,993

*Methods and systems for merging data during cache checking and write-back cycles for memory reads and writes*

Issued: May 6, 1997

Inventors: Richard P. Abato, et al

Assignee: IBM

Filed: April 29, 1996

Claims: 8

In response to memory reads or writes from a secondary processor, data is transferred into a buffer during a snoop cycle to a cache. The data in the buffer is merged with write-back data from the cache in a write operation. Data is provided directly from the buffer to the secondary processor and to main memory in a read operation.

#### 5,627,992

*Organization of an integrated cache unit for flexible usage in supporting microprocessor operations*

Issued: May 6, 1997

Inventor: Gigy Baror

Assignee: AMD

Filed: May 4, 1995

Claims: 32

A computer system having a cache that allows setting of caching policies on a page basis and a line basis. A status field for each cache block controls whether the cache-control unit operates in a write-through or copy-back write mode when a write hit access to the block occurs.

#### 5,627,981

*Software mechanism for accurately handling exceptions generated by instructions scheduled speculatively due to branch elimination*

Issued: May 6, 1997

Inventors: Michael C. Adler, et al

Assignee: Digital

Filed: July 1, 1994

Claims: 12

A method for generating an exception when a speculative instruction is committed. An exception flag, associated with the speculatively executed instruction, is tested at the commit point of the instruction. If it indicates an exception, an exception is raised at the commit point. □

**BUSES**

**FireWire getting hot.** IEEE 1394 is the "RCA jack" of tomorrow which, like its ubiquitous predecessor, will inspire a whole new wave of products and applications. Tom Cantrell, *Computer Design*, 10/97, p. 81, 4 pp.

**MIL-STD-1553 alternatives look to knock off the king.** Fibre channel is among the contenders with the specifications to offer a challenge. Duncan Young and John Wemekamp, DY 4 Systems; *Electronic Design*, 9/97, p. 162, 1 pp.

**The VMEbus picks up a flexible passenger.** If it's bridged properly, PCI lets designers easily add functionality. Jonathan Morris, Tundra Semiconductor; *Electronic Design*, 9/97, p. 166, 2 pp.

**Understanding and using the I<sup>2</sup>C bus.** This article describes the inter-IC control bus, a two-wire bus for providing a communication link between integrated circuits. James C. Flynn, *Embedded Systems Programming*, 11/97, p. 52, 10 pp.

**Ultra2 SCSI adds performance, but requires extra design effort.** New parameters and design challenges can be met by taking advantage of LVD signaling technology. Barry Caldwell and Larry Barnes, Symbios Logic; *Electronic Design*, 10/97, p. 71, 4 pp.

**IC DESIGN**

**VSIA pushes ahead on two new fronts.** Two key areas of VC-based chip design have moved quickly from development through the approval process: mixed signal and implementation/verification. Larry Waller, *Virtual Chip Design*, 11/97, p. 4, 3 pp.

**The system-on-a-chip: It's not just a dream anymore.** High-density processes, multiple memory technologies, and mixed-signal capabilities combine to realize what was once unattainable. Dave Bursky, *Electronic Design*, 10/97, p. 105, 7 pp.

**MISCELLANEOUS**

**Adapting Java for embedded systems.** Embedded tools vendors overcome Java's limitations to make it an attractive embedded development language. Peter Varhol, *Computer Design*, 10/97, p. 75, 4 pp.

**Tackle real-time applications with Windows NT.** Following the precedent of the PC architecture, the Windows NT operating system has entered the embedded-systems arena. For the right applications, it can be a contender. Richard A. Quinnell, *EDN*, 9/97, p. 61, 5 pp.

**Gas-gauge IC performs precise battery measurements.** Benchmarq's bq2018 improves battery management while reducing battery subsystem size and cost. Richard Nass, *Electronic Design*, 10/97, p. 39, 3 pp.

**PERIPHERALS**

**Seeing red: The IrDA protocol.** Infrared has a glowing future. It will play an increasingly important role in wireless data communications. John Canosa, *Embedded Systems Programming*, 11/97, p. 30, 12 pp.

**Choosing the right Ethernet switching chips.** Cost per port and firmware control are vital. Louis Pengue, PMC-Sierra; *Electronic Products*, 10/97, p. 45, 12 pp.

**PROCESSORS**

**RISC controller merges DSP and control functionality.** Tricore processors ease system design by combining the best features of DSPs and high-performance embedded controllers. Dave Bursky, *Electronic Design*, 9/97, p. 39, 4 pp.

**Enticing an infant market.** Semiconductor makers are aiming highly integrated processors at the nascent market for network computers. Tam Harbert, *Electronic Business Today*, 10/97, p. 49, 3 pp.

**PROGRAMMABLE LOGIC**

**Complex PLDs.** A directory of PLDs with at least 800 gate-equivalents. *Embedded Systems Programming*, Buyer's Guide 1997, p. 124, 3 pp.

**SYSTEM DESIGN**

**Sub-5-V circuits and ESD.** International electrostatic discharge standards pose challenges to portable systems designers. A new diode technology helps thwart ESD. Thomas Dugan, *Portable Design*, 9/97, p. 44, 2 pp.

**Flat-panel and CRT displays.** Here is a sampling of recently introduced flat-panel and CRT displays. *Electronic Products*, 9/97, p. 39, 7 pp.

**Power supply ICs.** Here are new and recently released ICs for use in power supply applications. *Electronic Products*, 9/97, p. 67, 9 pp.

**New paradigms ensure high-performance PCBs.** Major CAE companies are advocating that design engineers shoulder much of the burden of PCB analysis up front. Charles H. Small, *Computer Design*, 10/97, p. 60, 4 pp.

**Perspective on portable design.** Who says you can't take it with you? Mass storage for portables—both rotating disk and solid state—is expanding in density and variety. Terri Houston, *Portable Design*, 9/97, p. 31, 4 pp.

**DSP design issues for portables.** Signal processing of noisy dynamic signals places unique demands on designers. Here's an overview of some of the issues you'll likely face when using DSP. Chen Sagiv, DSP Group; *Portable Design*, 10/97, p. 44, 3 pp.

**Cool-running DSP design.** As switching speeds increase, power dissipation looms as a key DSP design factor. A clear advantage goes to the DSP chip that performs a function as quickly as possible, dissipating the least power. Mark Matson, Texas Instruments; *Portable Design*, 10/97, p. 49, 2 pp.

### ■ Death. Taxes. Comdex. In Portuguese.

Perhaps you've heard of it. It used to be that the Computer Dealers' Exposition (aka Comdex) came but twice a year. Now, with **Comdex/Miami**, visitors who miss the Las Vegas gala can still join the fun. The next installment arrives at the Miami Beach Convention Center December 9-11 and focuses on IT in Latin America. The four conference tracks include Business on the Net, the Business Desktop, Making Connections, and For Developers Only (with one day each on Java, ActiveX, and HTML). All Comdex/Miami presentations will have simultaneous translations into Spanish and Portuguese.

Prices range from \$25 for an exhibit pass to \$295 for an all-inclusive passport. For more information or to register, contact 617.433.1650 or visit [www.comdex.com](http://www.comdex.com).

### ■ New Compact Among CompactPCI Vendors

The PCI Industrial Computers Manufacturers' Group (with the dyslexic-unfriendly acronym of PICMG) has updated its collective specification for the **CompactPCI** standard. Revision 2.1 incorporates "additions and clarifications" over the previous version (1.0) released in 1995. Enhancements include new Eurocard mechanical specifications, consistent I/O connections, geographic addressing, and corrected interrupt routing.

Additional information is available at [www.picmg.org](http://www.picmg.org) or by calling PICMG (Wakefield, Mass.) at 781.246.9318.

### ■ Don't Drag Your Knuckles

This and other bits of social wisdom are to be found in *Interpersonal Skills for IS Professionals*, a seminar provided by Communication Workshop and the Center for Technical Communication. The seminar aims to take the "rough edges" off technical professionals who must deal with mere mortals.

More information is available from Communication Workshop (Port Washington, N.Y.) at 516.767.9590.

### ■ PC Chips Sales To Grow 15% per Year

So sayeth the prognosticators at In-Stat in their report, *Desktop PC Semiconductor Content and Integration Trends*. The report predicts that microprocessors will continue to snag the lion's share of chip revenue for the next five years, with a slight increase in the revenue share of memories.

Copies sell for \$2,995. For more information or to order a copy contact In-Stat (Scottsdale, Ariz.) at 602.483.4471 or set your browser to [www.instat.com](http://www.instat.com).

### ■ What's the Opcode for SEX?

This and other embarrassing questions are easily resolved on the **Microprocessor Instruction Set Cards** Web site. Located at [www.comlab.ox.ac.uk/archive/cards.html](http://www.comlab.ox.ac.uk/archive/cards.html), the site includes opcode references for dozens of microprocessors and microcontrollers of the past and present.

## MICRODESIGN RESOURCES

# Consulting and On-Site Seminar Services

MicroDesign Resources offers a full range of consulting and on-site seminar services to firms trying to find winning technology strategies in the microprocessor, 3D graphics, and personal computer industries.

Our analysts can:

- Identify emerging opportunities for growth
- Evaluate new technologies you are developing or adopting
- Analyze the competitive landscape and positioning challenges for your product
- Summarize your opportunities and challenges in the market

#### Scheduling:

Contact MicroDesign Resources at 707.824.4001 (toll-free 800.527.0288) or send e-mail to [cs@mdr.zd.com](mailto:cs@mdr.zd.com). We'll give you complete information on availability and costs.

## SUBSCRIPTION INFORMATION

To subscribe to *Microprocessor Report*, contact our customer service department by phone, 707.824.4001; fax, 707.823.0504; e-mail, [cs@mdr.zd.com](mailto:cs@mdr.zd.com); or Web, [www.MDRonline.com](http://www.MDRonline.com).

For European orders, contact Parkway Gordon by phone, 44.1491.875386, or fax, 44.1491.875524; or send e-mail to [parkway@rmpc.co.uk](mailto:parkway@rmpc.co.uk), or visit the Web at [www.parkway.co.uk](http://www.parkway.co.uk).

	U.S. and Canada*	Europe	Elsewhere
One year	\$595	£450	\$695
Two years	\$1,095	£795	\$1,295

\* Sales tax applies in the following states: GA, KY, MA, TX, and WA. GST tax applies in Canada.

*Microprocessor Report* (ISSN 0899-9341) is published every three weeks, 17 issues per year. Back issues are available on paper and CD-ROM. Volume reprints of individual articles are also available.

Ship to: