MICROPROCESSOR REPORT

THE INSIDERS' GUIDE TO MICROPROCESSOR HARDWARE

VOLUME 7 NUMBER 7 MAY 31, 1993

MIPS Reaches New Lows With R4200 Design

New Chip Can Power Sub-$2000 Systems with Pentium Performance

By Linley Gwennap

It will provide better price/performance than either a 486 or PowerPC 601, and yes, it will do Windows. The new R4200 design, revealed by MIPS Technologies, will deliver 55 SPECint92 and 30 SPECfp92, according to the company's simulations. The design, previously known as VRX or Ice, needs just 81 mm^2 of silicon, due in part to its unusual unified data path for integer and floating-point calculations. NEC, which will produce the chip, says the R4200 will cost 8000¥ (about $70) by mid-1994. The two partners hope that this low price will encourage system makers to use the chip in inexpensive Windows NT systems. The new chip will also be suitable for notebook systems and embedded applications.

Since MIPS, a subsidiary of Silicon Graphics, does not actually sell chips, the announcement was termed a "technology disclosure." The company expects to release the initial chip design for fabrication in June and receive first silicon soon thereafter. If the testing goes well, MIPS believes the initial design will be able to sample in 3Q93 and the production version will be ready by the end of the year. Although this schedule is aggressive, both the R4000 and R4400 sampled using first silicon with minor metal-layer fixes. NEC is more conservative, saying that production will ramp up during 1H94.

Despite its product number, the R4200 is a completely different design from the R4000/R4400. For example, superpipelining has been discarded in favor of a traditional five-stage pipeline. The new CPU maintains the full 64-bit MIPS-III architecture and is compatible with all current software. The first chips are expected to operate at 80 MHz internally with a 40-MHz system bus.

Unlike previous MIPS designs, the R4200 effort has been funded by only one of the semiconductor partners, NEC. Because of this arrangement, NEC has exclusive rights to produce the R4200 for one year. After that period, other partners may license the design from MIPS and NEC for an upfront fee proportional to the design costs. As a result, some of the six partners may choose not to market the R4200. MIPS expects that there will be multiple sources for the part in 1994, and hopes to announce a second source soon.

Designed for Low Cost

An important objective of the R4200 program was low cost. MIPS wants to move its architecture into the high-volume portion of the desktop market, where systems are priced between $1000 and $2000. This requires a chip priced to compete with the 486 chips used in the vast majority of these systems, which translated into two critical goals for the design team. First, the die should be as small as Intel's 486DX2 (81.5 mm^2 in a 0.8-micron process) to keep manufacturing cost down. Second, the chip must use a low-cost plastic package, which is much less expensive than the ceramic packages required by most RISC chips. The plastic package forced the designers to keep the power dissipation below 2 watts.

The die size proved to be the biggest challenge. Having access to NEC's 0.6-micron, three-layer-metal CMOS process provided a good start. Using four-transistor (4T) cells, the designers fit 24K of cache onto the small die with room for the rest of the CPU. Simulations showed performance in a Windows NT environment to be about 7% better with 16K of instruction cache and 8K of data cache than with the reverse configuration, so the team chose the larger instruction cache. Figure 1, a die plot, shows that the data cache could easily be expanded in a future, higher-performance design, if desired. In the current configuration, the die area beats its goal (the 486DX2's area) by 0.5 mm^2.

Integer and FP Data Path Combined

To further reduce the die area, the designers came up with an unusual strategy of performing floating-point calculations using the integer data path instead of having a separate FP unit. This doesn't mean that FP instructions are "emulated" using integer operations; the combined data path is enhanced to handle both types of math. Since the MIPS-III architecture specifies 64-bit integers, handling double-precision floating-point numbers in the same data path does not increase its width.

Figure 2 shows a simplified diagram of the data path. The 32 general (integer) registers and 32 FP registers are logically separate but physically reside in a single block of 64 registers. Since the R4200 issues one instruction per cycle, only two operands need to be fetched at a time. If the operands come from the floating-point registers, the 12-bit signed exponents are separated from the 53-bit signed fractions and sent to the exponent adder. (This example assumes double-precision FP data; the R4200 also handles single-precision FP values and 32-bit integers.)

The ALU on the left side of the figure handles integer calculations; fractions are simply treated as 53-bit integers. This ALU is capable of performing a full 64-bit add, shift, or boolean operation in a single clock cycle. It can also count the number of leading zeros in one cycle. The second ALU is a much smaller 12-bit adder that handles exponent calculations.

Integer instructions typically spend one cycle in the execution unit, but many floating-point operations take more than one cycle to complete. The summation registers and feedback loops allow values to circulate through the data path when performing these calculations. An FP add, for example, takes three cycles. During the first cycle, the fractions are added. In the next cycle, the fraction result is fed back into the ALU, which performs a leading-zero count. Finally, the fraction is sent back through the shifter for normalization while the exponent is adjusted appropriately. The resulting fraction and exponent are then combined in preparation for storing the result in the register file. (Note that for some FP adds, the fractions must be pre-aligned on the first cycle; in this case, the add occurs during the second cycle and a small normalization on the third cycle.)

A small block of logic examines the result of the exponent calculation to check for floating-point exceptions such as underflow and overflow. It also checks for conditions such as "equals one" or "equals zero." Any exceptions are signalled to the main control logic on the same cycle that the calculation completes.

This sequence of operations occurs in any floating-point unit; in a sense, the unified data path is an FPU that has been trained to do integer calculations as well. By simply extending the FP ALU and shifter to 64 bits, the separate integer ALU is eliminated. This not only saves die area but simplifies the control logic and reduces the number of bypass paths.

Table 1 lists the floating-point latencies for various operations on the R4200. The chip does not include dedicated hardware for either floating-point or integer multiplication, as these circuits are large and power-hungry. Instead, a three-bit Booth algorithm is used to accumulate the product.

By comparison, the R4400 uses the same 0.6-micron process as the R4200 but is designed for much better floating-point performance. The lower-cost chip suffers from the lack of a dedicated 64-bit multiplier and the inability to pipeline multiplies, a significant part of the SPECfp92 suite. FP square root and integer divide take about the same amount of time on both chips; with its longer clock period, the R4200 can calculate twice as many bits per cycle as the R4400.

The R4200 compares quite well to the 486DX2, as shown in Table 1, particularly on FP addition. One drawback of the R4200's combined data path is that long-

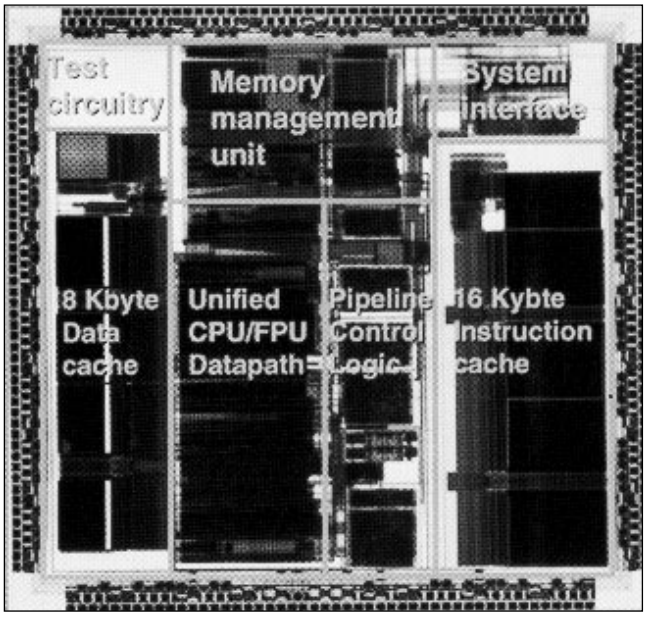

Figure 1. A die plot of the R4200, which will be 8.8 mm x 9.2 mm (346 mils × 362 mils). It uses 1.3 million transistors.

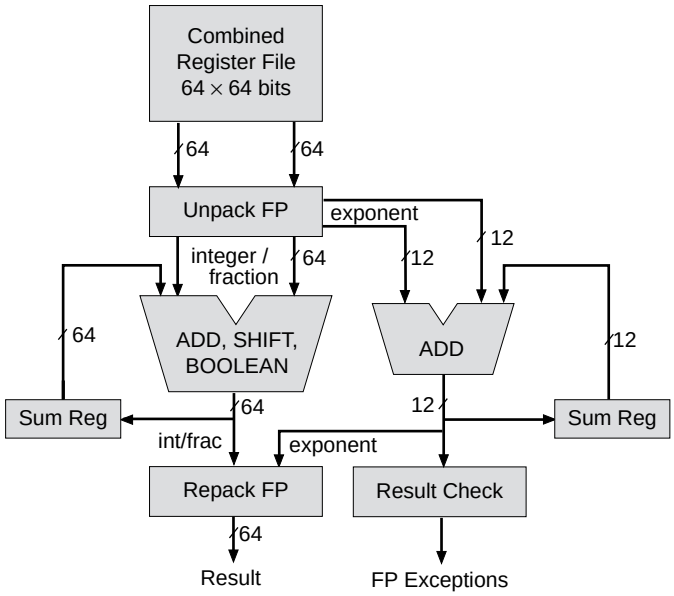

Figure 2. Simplified diagram of the R4200's unified integer/FP data path. Data from immediate values and operand bypassing are not shown.

latency FP operations tie up the entire CPU; many other processors can continue with integer instructions while the FPU performs a multiply or divide.

Pipeline Similar to R3000

The R4200 reverts to the five-stage pipeline used by the R3000; the R4400's eight-stage pipeline was deemed too complex for a low-cost chip. The five-stage pipe requires less bypass circuitry and has only a single-cycle stall for branches and loads. With fewer and shorter stalls, the shorter pipeline can use a lower clock rate to achieve a given performance level, thus reducing power consumption.

The R4200 pipeline is very similar to the R3000 implementation except that the new chip uses a virtually-indexed cache, allowing it to perform the address translation in parallel with the cache access.

The separate caches are direct-mapped and physically tagged. The instruction cache uses 32-byte lines while the data cache uses 16-byte lines. The data cache uses a write-back protocol and allocates lines on writes. A one-entry write buffer prevents the CPU from stalling on writes to memory or I/O, an improvement over the R4400, which only buffers I/O writes and graphics data.

Unlike previous MIPS processors, the R4200 does not support instruction streaming; after an instruction miss, the CPU waits for the entire line to be filled before resuming. On a data miss, however, the system interface returns the critical doubleword first, and the CPU restarts as soon as it receives that data; this feature is relatively trivial, since there are only two doublewords per data cache line.

The memory-management unit (MMU) on the new chip is essentially identical to the R4000's MMU except that the number of TLB entries is reduced from 48 to 32 to save die area. Each entry can map either a standard 4K page or a 16M page. Like previous MIPS designs, the R4200 has a separate two-entry instruction TLB ("micro-TLB") to allow simultaneous translation of instruction and data references in most situations.

To reduce cost and complexity, the R4200 restricts many of the selectable features of the R4000. For example, the cache line size, page size, and system-bus clock and write data pattern can all be set to a wide range of values in the R4000. The newer chip restricts these parameters to one or two settings each. Unlike the R4000PC, which simply doesn't bond out the secondary-cache bus and multiprocessor signals, the R4200 has removed all control signals, logic, and instructions that refer to the external cache or involve MP support.

Power Management Emphasized

Another key objective for the R4200 design was low power, due in part to the need for a plastic package but also to allow the chip to be used in notebook computers and other portable systems. It is the first workstation-compatible RISC processor to incorporate a wide range of power-saving features. As a result, the chip is rated at just 1.5 W for typical operation.

	R4200 (80 MHz)		R4400 (150 MHz)		486DX2 (66 MHz)	
	cycles SP/DP	time (DP)	cycles SP/DP	time (DP)	cycles SP/DP	time (DP)
FP ADD	3/3	38 ns	3/3	20 ns	8–20	150 ns
FP MUL	11/20	250 ns	7/8	53 ns	11/14	210 ns
FP DIV	29/58	725 ns	23/36	240 ns	73/73	1095 ns
FP SQRT	30/59	738 ns	54/112	747 ns	87/87	1305 ns
INT MUL	13/24	300 ns	10/20	133 ns	26/—	780 ns
INT DIV	37/69	863 ns	69/133	887 ns	43/—	1290 ns
SPECfp92	30 (simulated)		90 (Indigo2)		16 (Deskpro)	

Table 1. Maximum latencies for various math operations on MIPS and Intel processors. Note that the R4400 can overlap two FP multiplies. (Source: vendor data books)

The design uses several familiar techniques to minimize power usage. It operates at 3.3V, which offers a 50% power reduction from a 5V design. Functional blocks are powered down on a cycle-by-cycle basis when not in use. Execution units that are not performing necessary calculations are fed a constant input to prevent switching, while the write-back cache reduces system bus activity and thus system power.

Other power savings come in less common ways. The caches are divided into four banks each, and the bank select is decoded early so that only the requested bank is cycled; power-hungry sense amps and logic for the other banks are not used. The early decoding slightly increases the cache access time, but this is not a critical path in the R4200.

The instruction cache delivers two instructions (64 bits) at once and thus is often powered on only every other cycle. The micro-TLB holds the two most-recent instruction translations; the main TLB is not powered if the micro-TLB hits.

The R4200 does not use a static design. MIPS claims that a dynamic design uses less die area and less

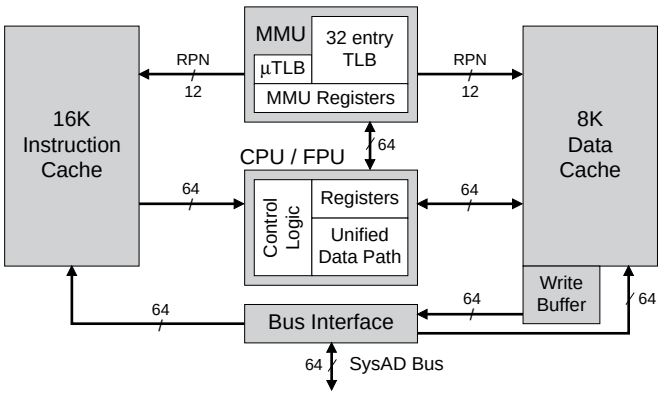

Figure 3. Block diagram of the R4200 shows 64-bit paths for all data except real page numbers (RPNs) from the TLB.

Price and Availability

The R4200 is not yet announced as a product. The chip will initially be produced by NEC, although MIPS expects other sources to become available. NEC expects to begin sampling the R4200 in 3Q93 with production in 1H94. NEC's production pricing target is about $70.

NEC, M/S MV4580, 475 Ellis Street or PO Box 7241, Mt. View, CA 94039; 800/366-9782 or 415/960-6000, fax 800/729-9288.

power when operating at full speed, and believes there is little customer demand for being able to stop the clock without removing power. If the R4200 is shut down, all internal caches and registers must be saved before removing power. The chip supports a quarter-speed (20 MHz) operating mode, enabled by software, that reduces power by nearly 75%. There is no system-management interrupt, as in many x86 processors, because Windows NT uses a different method for power management.

R4000 System Interface Maintained

MIPS has attempted to retain as much of the R4000 interface as possible to simplify the system design task. The R4200PC is available in the same 179-pin PGA package as the R4000PC and uses the same 64-bit multiplexed address/data bus. There are a few changes, however, that prevent the new chip from being simply dropped into an R4000PC system design.

The biggest difference is that the R4200 uses a 3.3V supply instead of 5V. All of the components that connect to the CPU must operate at 3.3V signal levels as well. Although this may be inconvenient for system designers, the R4200's 0.6-micron process is optimized for lower-voltage operation, and voltage-translation buffers would require valuable die area and power.

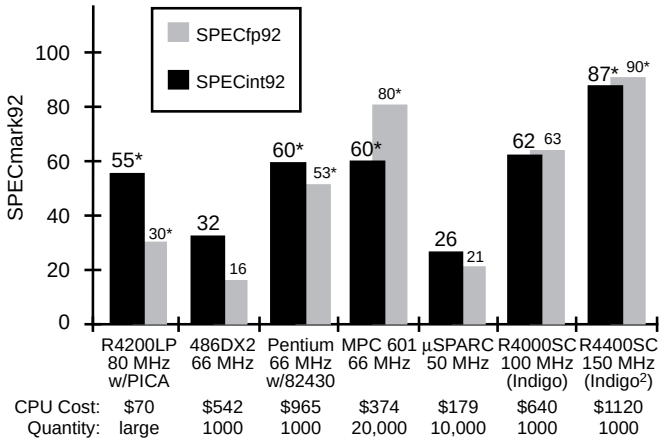

Figure 4. Simulated SPEC92 performance of the R4200 (with external cache) compared to reported performance of other announced processors. (Source: SPEC except *vendor estimates)

Another change is that the processor initialization process has been greatly simplified. The R4000 uses a complex scan-based initialization scheme to set up all of the various configurable options. Since the R4200 has many fewer options, it can be initialized simply by asserting RESET*.

Three newly-defined pins select the critical options. One chooses big- or little-endian operation. Another sets the bus write pattern to either DDx (two doublewords every three system-bus cycles) or Dxx (one per three cycles); these are the only two data patterns supported for writes. A third pin can disable the clock-doubler (PLL); this requires the system to generate an 80-MHz clock but allows it to quickly drop the CPU frequency to a much lower rate when the system is idle. All other CPU options are selected by software-writable registers.

A few R4000 system designs will need additional modifications to use an R4200. It is possible to design an R4000 system with a four-word instruction-cache refill, while the new chip demands eight instruction words at a time. The R4200 is also less flexible regarding the timing of the bus clock signals (TClock and RClock) than the R4000. MIPS believes that few designers of low-cost systems will be affected by these changes.

Unlike the R4000, the R4200 will also be available in a 208-pin PQFP. This version, dubbed the R4200LP, is expected to be more popular, as the plastic package is smaller and less expensive. Thermal issues limit the LP version to 80 MHz; speedier versions may eventually be available in the ceramic PGA.

Since the R4200 does not directly control a secondary cache, MIPS expects that many system vendors will use an external cache controller. With the same system interface as the R4000, the new CPU will be compatible with various low-cost chip sets available for this purpose *(see **070501.PDF**)*. These chip sets interface to standard PC memory and peripherals, enabling system makers to use the R4200 at PC price points.

R4200 Outguns 486 By Two-to-One

Intel's 486DX2 is the prime target of the R4200. The DX2 will be the mainstay of the Windows NT market for the next year or two, and MIPS hopes to grab a piece of that market by offering nearly twice the performance at a lower price. Figure 4 shows the price/performance gap. The R4200 is expected to achieve nearly the same integer performance as Pentium, which requires over three times the die area and is much more expensive.

Another key competitor, both in the workstation arena and possibly for Windows NT, is the PowerPC 601. The 601 offers slightly better integer performance but uses a 50% larger die and costs far more at $374. To be fair, we must note that the R4200 price is for 2Q94, but it seems unlikely that either Intel or Motorola would drop their prices by 80% over the next year.

Other low-end competitors include Sun/TI's microSPARC and two unannounced chips, HP's PA7100LC and DEC's 21066. All of these chips integrate memory and system interfaces but use more than twice the die area of the R4200. System logic for the MIPS processor costs $149 today (Acer's PICA chip set) and may fall below $100 by the end of the year, making it unlikely that the HP or DEC chips will be less expensive than the combination of an R4200 and a system-logic chip set.

MicroSPARC, which sells for a competitive $179 despite its huge die size, has much lower performance than the R4200, although a planned 75-MHz upgrade will close the gap somewhat. The Sun design uses SBus for its system interface, leaving it unable to access low-cost PC expansion cards. There are also no PC operating systems available for SPARC, although SunSoft is developing a "WABI" to allow Windows applications to run under its Solaris OS *(see **0707ED.PDF**)*.

Perhaps the most potent competitor for the R4200 will be IDT's Orion (see sidebar). The two products were thought to have the same objectives, but recently both MIPS and IDT have tried to position Orion at a higher price and performance level than the R4200, allowing the two to co-exist in the market. The performance gap between the two designs is relatively small, however, and their manufacturing costs should be similar, so the two products may overlap significantly. MIPS will be happy if either chip does well.

The achilles' heel of the R4200 is its relatively low floating-point performance. Although it matches up well with the 486 and microSPARC, the 601 and other chips greatly outperform it. This reduced emphasis on FP indicates a clear focus on NT users, whom MIPS expects will rarely need floating point; those with technical applications may need an R4000 (or Orion) system.

Ready to Battle the Intel Giant?

The R4200 should quickly obsolete the R4000PC and makes the R4000SC unattractive unless floating-point performance is critical. It will enable excellent UNIX notebook computers, but the market for such systems is small. Its low cost and emphasis on integer performance make it well-suited for embedded applications such as printers, X-terminals, and network routers. The layout is designed so the semiconductor partners can easily modify the cache sizes and remove the MMU to further customize the design for specific applications.

Windows NT is the most attractive target market, with a long-term potential of millions of units per year. The R4200 will have a huge price/performance advantage over Intel's processors at a price low enough to attract PC vendors. Microsoft expects to deliver NT-on-MIPS concurrent with the x86 version (see page 5), and inexpensive system-logic chip sets are already available. All that's missing is the system vendors; so far, Acer is the only major PC maker strongly committed to the MIPS architecture. MIPS has delivered the battle plan, but it remains to be seen if there are enough troops willing to go against Intel. ♦

Orion Has Similar Goals

Integrated Device Technology (IDT) is sponsoring QED, a design firm, to create a low-cost MIPS processor code-named "Orion" *(see **061507.PDF**)*. The Orion project began with similar goals as the R4200: a low-cost MIPS CPU with Pentium-class performance. Based on the R4200 announcement, it now appears that the QED chip will provide about 30% better integer performance due to its 100-MHz clock rate and 16K data cache. Orion also has a full floating-point unit that should offer much better FP performance than the R4200. IDT would not disclose pricing for its unannounced processor but believes that Orion will be priced somewhat higher than the R4200.

This price differential is based on performance and not manufacturing cost; IDT reports that Orion is still on target to have a smaller die size than the R4200, despite its larger data cache and separate FPU. QED expects to release the initial design for fabrication in June and hopes to be neck-and-neck with the R4200 in the race to full production. One advantage for the R4200 is that its power usage is about half of Orion's, making the MIPS design better for portable systems. For other applications, Orion may offer better performance at a potentially lower cost—if it meets its design goals.

