620 Fills Out PowerPC Product Line

New 64-Bit Processor Aimed at Servers, High-End Desktops

by Linley Gwennap

At the recent Microprocessor Forum, Motorola and IBM revealed the high-end PowerPC 620, the last of the four chips from the original 1991 roadmap. The new processor is the first in a line of PowerPC chips intended for servers and high-performance workstations.

Optimized for performance instead of price, the 620 will achieve 225 SPECint92 and 300 SPECfp92 at 133 MHz, according to Chin-Cheng Kau of the Somerset design center. This integer performance is 40% better than for the 100-MHz 604, while the FP score is nearly doubled. To improve server performance, the design includes 64K of on-chip cache, a direct interface to a large external cache, a 128-bit bus to main memory, and a glueless multiprocessor interface, all firsts for PowerPC. The 620 is also the first PowerPC chip to implement the full 64-bit version of the architecture.

Pushing the performance pedal to the floor, however, leaves the chip redlined on cost. The new features all add significant cost to either the package or the die, both of which set or approach records for general-purpose microprocessors. At $380, the estimated manufacturing cost of the 620 is three times that of the 604, although it is in the same range as other next-generation RISC processors. Neither company has announced a price for the new processor, but we expect it to exceed $1,200 when the 620 debuts in 2H95.

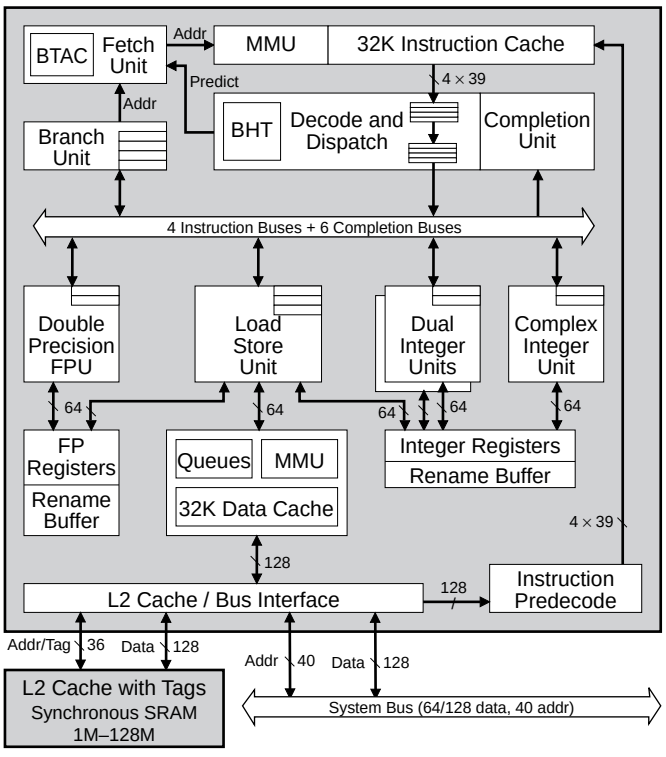


Figure 1. The PowerPC 620 block diagram is similar to the 604's, but the new chip adds a predecode unit and a level-two cache controller.

PowerPC Extended to 64 Bits

The 620 is the first implementation of the 64-bit version of the PowerPC architecture *(see **070302.PDF**)*. Like other 64-bit architectures, it includes 64-bit integer and floating-point registers. It extends the PowerPC segmented addressing scheme to allow 2^{36} segments of 2^{28} bytes (256M) each; using the proper segment tables, these can be arranged as a 64-bit linear address space. A segment lookaside buffer caches 20 of these table entries to speed translation.

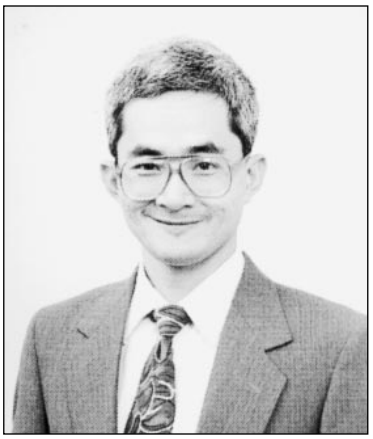
The 620 also includes the 16 segment registers required for compatibility with 32-bit PowerPC programs, which use 16 segments of 2^{28} bytes each. These segment registers are 64 bits wide in the 620, allowing an OS to retain the current segment model while accessing a total virtual address space of one heptabyte (2^{80} bytes).

It will take some time for both 64-bit applications and operating systems to become available for the 620. The processor supports mixed environments; for example, a 64-bit operating system can run 32-bit applications simply by setting a mode bit in the CPU before calling the application. (Care must be taken in passing parameters in this situation.) A 32-bit operating system can run 64-bit applications by reversing this procedure; in this case, the application could perform 64-bit integer arithmetic but could not access the 64-bit address space.

Branching Enhanced for Business Apps

The 620 microarchitecture is quite similar to the 604's *(see **080501.PDF**)*, although the two chips were developed separately at the Somerset design center. Like its predecessor, the 620 can issue up to four instructions per cycle and has dual units for common integer instructions, along with a third unit for complex integer instructions (e.g., multiply and divide). As Figure 1 shows, the design also includes a load/store unit, branch unit, and floating-point unit.

Like the 604, the 620 uses dynamic branch prediction based on a branch history table (BHT) and branch target address cache (BTAC). These structures



CLARENCE TOWERS

IBM's Chin-Cheng Kau describes the microarchitecture of the new PowerPC 620.

allow a taken branch to be processed without delay if it hits in the BTAC or with a single-cycle penalty if it does not. To improve the hit ratio, the 620 quadruples the size of the BTAC, which now contains 256 entries. The BHT is also quadrupled, to 2,048 two-bit entries, boosting branch prediction accuracy to 90% on SPECint92.

This rate is better than that of other processors due to PowerPC's condition-code architecture. The compiler strives to place the instruction that generates a condition code well ahead of the conditional branch. If the two instructions are separated by two cycles, the 620 can "predict" the branch by examining the calculated condition code, improving the prediction accuracy.

The new design eliminates the extra cycle needed by the 604 to group and dispatch instructions. Unlike other next-generation RISC chips, the 620 uses a traditional five-stage pipeline, reducing the mispredicted branch penalty from three cycles to two. To achieve this result, the designers added a predecode unit that calculates seven extra bits per instruction, which are stored in the instruction cache. This extra information speeds the decode and dispatch processes, allowing both operations to occur in the same pipeline stage.

The 620 retains the separate four-entry instruction and dispatch buffers of the 604. Instructions flow into the dispatch buffer only in pairs, so in some situations only three instructions are available for dispatch instead of the maximum of four.

All of these improvements from the 604 will particularly improve the 620's performance on business applications, which typically contain more frequent and less predictable branches than technical code. The larger BTAC and BHT will keep the pipeline flowing smoothly during frequent predictable branches. Even for branches that are difficult to predict (for example, "if status=married, then…"), the shorter pipeline keeps delays to a minimum.

Improved Out-of-Order Execution

Although other RISC vendors are just beginning to implement out-of-order execution, the 620 is the third PowerPC chip to use this technique. Like its predecessors, the new design includes "reservation stations" in each of the function units to hold instructions that are waiting for operands. The integer and floating-point units each have two reservations stations, as in the 604; the new branch unit has four reservation stations, permitting speculative execution beyond as many as four conditional branches.

CLARENCE TOWERS

Brad Beavers of Motorola discusses the 620's high-bandwidth system interface.

The 620 improves upon its predecessor in the load/store unit by adding a third reservation station. Furthermore, this unit now allows loads and stores to complete out of order. The completion unit keeps track of the low-order 12 bits of the effective address for loads and stores; if a store address matches the address of a load that was executed out of order, the instruction pipeline is flushed and the load is re-executed. Although flushing the pipeline causes a serious performance hit, this situation should occur very infrequently for most programs.

The load/store unit will also perform some loads speculatively, that is, when a preceding branch instruction has been predicted but not resolved. To prevent operations that cannot be undone, the 620 will not speculatively execute stores of any kind or loads from non-cachable (I/O) segments. Because most loads and stores are queued, as in the 604, these operations typically will not stall the processor pipeline.

The 620 handles unaligned loads and stores in hardware in both big- and little-endian modes, allowing it to do well in mixed environments with older Apple Macintoshes and x86-based PCs.

The other function units are quite similar to those in the 604. The integer units, of course, are expanded to 64 bits in width. The latency of floating-point loads is reduced from three cycles to two, and the latency of floating-point divides is improved to 18 cycles. The 620 adds a hardware square-root function that completes in 22 cycles; in the 604, square roots are calculated in software. These changes will improve performance on Spice circuit models, among other applications. As in the 604, all FP operations are performed in double-precision; there is no speed improvement for single-precision calculations.

Two-Level Cache Structure Added

Although the 60x PowerPC chips can operate with second-level caches, all control functions for the L2 cache must be implemented in external logic. Furthermore,

	PPC 620	PPC 604	PPC 601	PPC 603
Max Clock Rate	133 MHz	100 MHz	100 MHz	80 MHz
Total Cache	64K	32K	32K	16K
Peak Issue Rate	4 instrs	4 instrs	3 instrs (2+branch)	3 effective (2+branch)
Load/Store Unit?	Yes	Yes	No	Yes
Dual Integer ALUs?	Yes	Yes	No	No
Reg. Renaming?	Yes	Yes	No	Yes
Branch Prediction	Dynamic	Dynamic	Static	Static
Supply Voltage	3.3 V	3.3 V	2.5 V	3.3 V
Maximum Power	30 W	13 W	4 W	3 W
CPU Clock Input	1×…1/4×	1×…1/3×	2×	1×…1/4×
Package	BGA-625	CQFP-304	CQFP-304	CQFP-240
IC Process (drawn)	0.5 µm 4 metal	0.65 µm 4 metal	0.5 µm 5M + local	0.65 µm 4 metal
Transistors	6.9 million	3.6 million	2.8 million	1.6 million
Die Size	311 mm^2	196 mm^2	74 mm^2	85 mm^2
Est. Mfg. Cost*	$380	$125	$65	$45
Availability	3Q95	4Q94	4Q94	3Q94
Estimated SPEC92	225 int	160 int	105 int	75 int
@ Max Clock Rate	300 fp	165 fp	125 fp	85 fp

Table 1. The PowerPC family now spans a wide range of cost and performance, from the low-end 603 to the new 620. (Source: Somerset except *MDR Cost Model)

these chips share a single interface between the L2 cache and other system accesses. The 620 solves both of these problems, significantly improving performance.

The first level of caching is implemented by the 32K instruction cache and 32K data cache, both on the chip. These caches are twice as large as those in the 604 and are eight-way set associative. Increasing both the size and the associativity from the 604 will boost the cache hit rate, particularly for larger applications. The caches are otherwise similar to those in the 604.

The 620 uses a new technique to achieve "effective" eight-way associativity. A traditional eight-way cache delivers eight times the amount of data as a direct-mapped cache; a set of multiplexers then selects the correct data, as indicated by the tag-comparison result. The 620, in contrast, integrates a content-addressable memory (CAM) into the front end of the cache access.

The CAM determines the correct set before the array is accessed, eliminating the wide buses and multiplexers of the traditional design as well as the time needed to select the correct data. This technique also reduces power consumption for the large cache. Simulations show that this design has about the same hit rate as a traditional eight-way cache, although performance is slightly different, depending on the data set.

The integrated L2 cache controller supports external caches that range in size from 1M to 128M. The unified L2 cache is direct-mapped and protected by ECC. It is accessed through a dedicated 128-bit data bus, refilling an on-chip cache line in four accesses. The cache data and tags both use synchronous SRAMs with CMOS or GTL interfaces.

The external cache can operate at the same speed as the CPU or, to reduce system cost, at one-half or one-third of the CPU clock rate. IBM and Motorola expect most systems to use a half-speed cache; the performance estimates assume this configuration. At this speed, the penalty for a load that hits in the L2 cache is eight cycles (8-2-2-2 access rate). With a full-speed cache, this is reduced to 6-1-1-1, a 36% improvement. Note that instruction refills take an extra cycle due to the predecoder. According to Somerset's Brad Beavers, a full speed cache will improve SPECint92 performance by about 5%.

Moving the L2 cache controller on chip allows a dedicated cache interface, removing cache accesses from the system bus. The 620's wide interface will deliver at least twice the bandwidth of the 60x system bus, or four times with a full-speed cache. The separate cache bus frees the 620's system bus for memory and I/O traffic and allows the two buses to run at different clock speeds, if desired.

The MMU structure varies significantly from previous PowerPC chips, which perform a two-step address translation: the 32-bit effective address is converted to a 52-bit virtual address (24-bit segment plus 28-bit offset), which is then translated into a 32-bit physical address. The 620 contains dual 64-bit TLBs that translate effective addresses directly to physical addresses. Segment translations are downloaded from a unified 128-entry TLB whenever there is a context switch. The new design speeds the translation process.

Memory Bandwidth Increased

The 620 system bus permits a low-cost 64-bit interface or a high-performance 128-bit design. Even the 64-bit version will provide much better memory bandwidth than the 64-bit 60x bus, as all cache traffic is shifted to the dedicated cache bus. This mode should be adequate for uniprocessor servers or multiprocessor workstations. The 128-bit mode may be required to meet the high bandwidth requirements of MP servers.

Up to four 620 chips can be connected in an MP configuration without any glue logic; all simply share the same system bus. The processors handle all cache consistency checking; the on-chip cache tags are dual-ported, so there is no overhead for snooping. (L2 snoops block that cache for one L2 cycle.) The bus uses a split-transaction protocol to overlap requests from multiple processors; one instruction fetch, one data load, and up to three stores can be outstanding at any time from each CPU on the bus.

The 620's system bus uses 3.3-V signal levels to support operation at up to 75 MHz and possibly higher in future implementations. The bus can operate at one-half, one-third, or one-quarter of the CPU speed. The address bus is 40 bits wide, allowing physical memory sizes of up to 1,024 gigabytes. Both the address and data buses are parity protected.

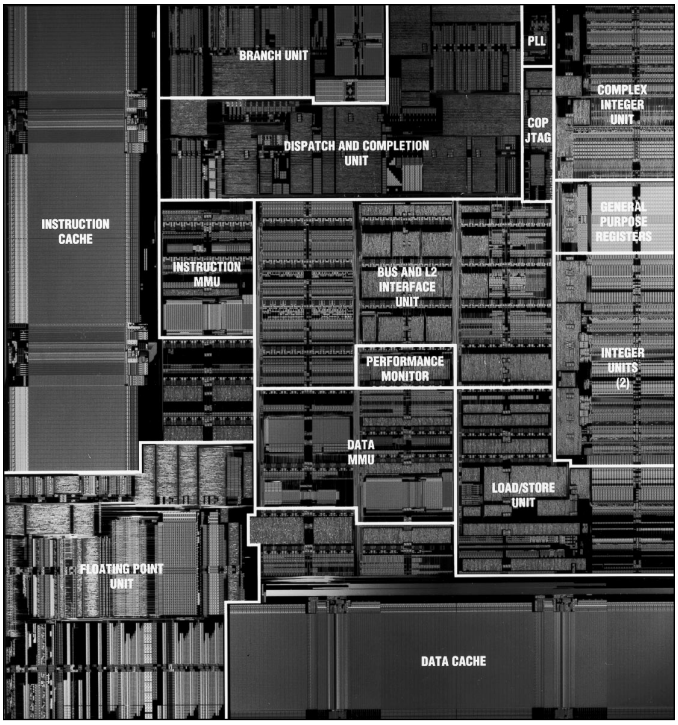


Figure 2. The 620 die measures 17.1 mm × 18.2 mm in a 0.5-micron four-layer-metal CMOS process and contains 6.88 million transistors. The chip uses area bonding and has no pad ring.

Cost of Complexity Is High

The 620 uses a slightly different manufacturing process than the 603 and 604: a four-layer-metal version of a process that IBM calls CMOS-5S *(see **080504.PDF**)*. This process reduces the gate oxide thickness, speeding the transistors. The faster transistors, along with improved circuit designs, deliver a clock rate 33% faster than that of the 604.

The added features increase the die size of the 620 to 311 mm^2, 59% larger than the 604. The new chip would have been even larger, but Motorola agreed to adopt IBM's C4 (flip-chip) bonding method *(see **071304.PDF**)*. This method eliminates the need for a pad ring, as Figure 2 shows. The major contributors to the area increase are:

- Larger caches (including predecode bits) .....65 mm^2
- Wider bus and new L2 cache controller .....30 mm^2
- 64-bit integer registers, ALUs, MMUs .....15 mm^2
- Larger FPU (improved FP performance) .....15 mm^2
- Improved branching (BTAC, BHT).....5 mm^2
- Eliminated pad ring .....–15 mm^2

The dispatch/completion and load/store units remain about the same size as in the 604.

The 620 design uses 6.88 million transistors, more than two-thirds of them in the large caches. Roughly 2.2 million are required for the CPU core and interfaces, compared with 1.8 million for the higher-performance 21164. Fast scalar CPUs need less than one million transistors, indicating the complexity of the 620 design.

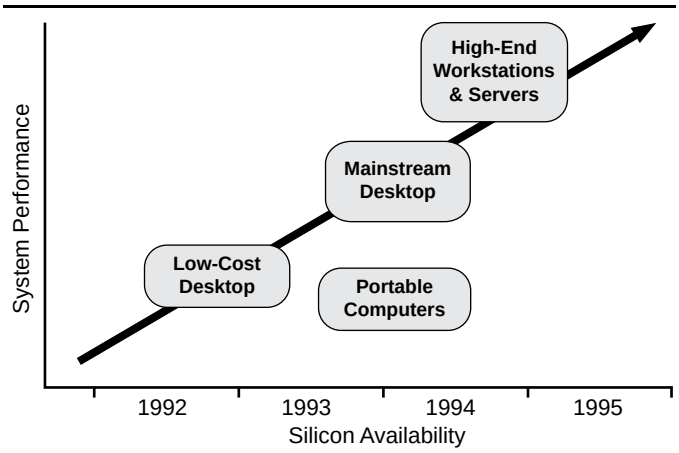


Figure 3. The original PowerPC roadmap published in 1991. The four boxes became the 601 (low cost), 603 (portable), 604 (mainstream), and 620 (high end).

The two 128-bit interfaces, along with miscellaneous signals and power, require a 625-pin package. A PGA of this size would be quite expensive, so the designers opted for a BGA instead, which eases the cost by about $25. Still, this package is expensive and has the highest pin count of any merchant microprocessor. The MPR Cost Model *(see **081203.PDF**)* estimates that the total manufacturing cost of the 620 will be about $380, just under the $430 mark set by the 21164.

The chip is no slouch in the power department either: the companies estimate that the 620 will dissipate about 30 W (maximum) at 133 MHz, a figure that is becoming standard for next-generation CPUs.

PowerPC Completes Original Plan

Figure 3 shows the original PowerPC roadmap announced in October 1991 at the Microprocessor Forum. The size of the boxes and the term "silicon availability" leave this chart open to a certain degree of interpretation, but IBM and Motorola have generally indicated that the midpoint of each box is supposed to represent the availability of first samples to the partners (not necessarily general sampling).

By this interpretation, the companies have done an admirable job of meeting their original plan. The 601 and 603 are spot on schedule, with the 604 and 620 lagging by a couple of months. Unlike many other technology partnerships *(see **0813MSB.PDF**)*, this joint effort has thrived.

Having established a good track record, the vendors are uncharacteristically mum about future plans. In contrast with Intel, which designs high-end chips and moves them downward in price over time, the PowerPC team plans to keep its processors aimed at specific market segments, tuning them for the needs of each. In the immediate future, this means die shrinks and minor enhancements to the announced designs; in the longer term, a new set of cores will eventually debut.

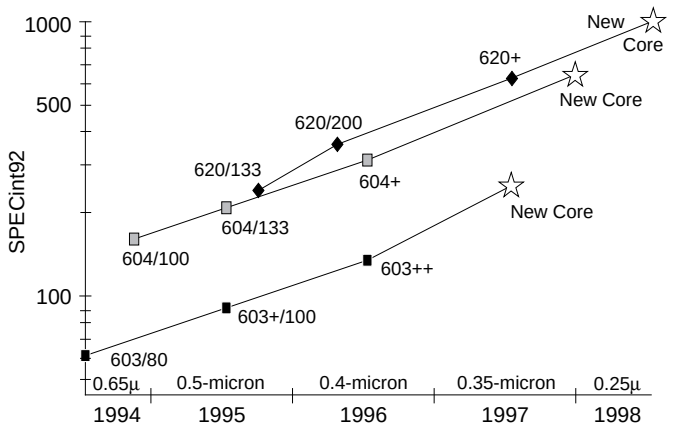


Figure 4. Projected roadmap of future PowerPC processors shows continued growth along three main product lines. (Source: MPR)

Figure 4 shows our projection of future PowerPC processors. In 1995, both the 603 and 604 will shrink to 0.5-micron CMOS, increasing clock speeds. The 603 is likely gain a larger cache (603+); such a part would obsolete the 601. In 1996, the processors will move to a new process, perhaps 0.4-micron, probably with some design tweaks to take better advantage of higher clock rates. New cores should debut in 1997–98, starting at the low end and working upward.

Based on SPECint92, the 620 does not appear to offer a significant benefit over the 604, particularly if the 0.5-micron 604 reaches 133 MHz. The key advantage of the 620 is its higher memory bandwidth for large technical and commercial applications, which is not apparent from the relatively small SPECint92 benchmarks.

One wildcard is the widely rumored PowerPC 615, a processor that includes hardware acceleration for x86 emulation *(see **080704.PDF**)*. Sources indicate that the 615 is nearing tape out and could be shipping in systems in 4Q95. Both IBM and Motorola will apparently source the part. If the acceleration features of the 615 are accepted by the market, these features are likely to be incorporated into other PowerPC processors, but it isn't clear how soon and to what degree this will happen.

PowerPC Performance Lags Other RISCs

While the PowerPC line was building to a climax, Motorola and IBM said that the first chips, for a variety of reasons, did not show the full performance potential of the PowerPC architecture. The 620 is supposed to be the chip that moves PowerPC into the rarified echelon of Alpha and PA-RISC performance. From initial appearances, however, it fails in this task.

Digital's 300-MHz 21164 *(see **081201.PDF**)* exceeds the performance of the 133-MHz 620 by 50%, according to estimates from the two companies. Perhaps the 620 estimates are a bit conservative, but not by enough to close such a large gap. Furthermore, the 620 is expected to reach the market 6–9 months after the 21164; in fact, Digital may push its chip even faster by the time the 620 is available.

In 2H95, Sun, MIPS, and HP all expect to deliver new processors with significantly better performance than the projections for the 620. Granted, the PowerPC vendors have a better track record of delivering what they say, but one or more of these other vendors is likely to approach Digital's performance. It remains to be seen how all these chips will stack up in real systems.

Although the 620 excels in "efficiency" (measured in SPECint92/MHz), it is hampered by the lowest clock rate among the next-generation processors. The overly complex Brainiac design style *(see **0703ED.PDF**)* of the PowerPC chip leaves it trailing the simpler Speed Demons in performance. We expect that the 620 will do little to boost the sales of IBM's high-end workstations.

From the PowerPC point of view, the major competition is Intel, not the ragamuffin RISC vendors, so perhaps this performance deficit is not important. The new chip is clearly far ahead of Pentium in performance, but by the time the 620 is available in systems, Intel's P6 will be imminent. Based on what little is known about the unannounced P6, it appears that the 620 is somewhat faster than the x86 chip on SPECint92 but certainly nowhere near the 2× that the PowerPC backers had led us to expect.

As a CPU for high-performance servers, the 620 should fare better. The processor has excellent memory bandwidth, large on-chip caches, and support for low-cost multiprocessor configurations. IBM and Motorola claim that the chip will outperform its rivals by a factor of two on metrics such as transactions per second, but without working systems, it is impossible to make such comparisons.

In desktop systems, PowerPC has demonstrated that it can deliver the same performance as x86 chips for half the CPU price. It would be more interesting, however, if it could deliver twice the performance for the same price. So far, only Alpha has been able to maintain its performance at twice the level of the fastest x86 processors. The 620 will do well as a server engine, but it appears that it will fail to open a significant performance gap against P6. ♦

Price & Availability

The PowerPC 620 will be marketed by IBM and Motorola. Neither company has announced price or availability for this part, but they expect samples to be ready in 2Q95, with production shipments in 2H95. For more information, call Motorola at 800.845.MOTO or IBM Microelectronics at 800.POWERPC, or contact your local IBM or Motorola sales office.

