

MICROPROCESSOR REPORT

THE INSIDERS' GUIDE TO MICROPROCESSOR HARDWARE

Exponential's PowerPC Blazes

At 533 MHz, Bipolar x704 Smokes Competition But Burns 85 W



by Linley Gwennap

There's a bright new star in the PowerPC sky: at last week's Microprocessor Forum, Exponential Technology debuted its first product, the x704, which the company expects to outrun all PowerPC processors when it ships next spring. This performance is achieved with a bipolar design, a technique most microprocessor vendors gave up on years ago. The bipolar circuits allow the chip to reach speeds in excess of 500 MHz, enabling high performance from a CPU of modest complexity. The x704 combines this bipolar logic with CMOS memory to improve density and cost.

The downside is a voracious appetite for power: the x704 burns as much electricity as a fast PowerPC 604 chip and a 60-W light bulb combined. Although cooling the x704 requires an extra fan and a large heat sink, it can be done with minimal changes to a standard desktop system. With these changes, a system maker can gain a performance increase of 50% over the fastest PowerPC chips currently available. This gap will narrow, however, by the time the x704 ships.

With this strong performance, the startup is targeting the high end of the Macintosh market as well as high-performance Windows NT systems. In the latter segment, the x704 will join Digital's 21164 as the only NT processors that offer more than 10 SPECint95. In the Mac market, the x704 should have a narrow but clear performance lead when it debuts.

The Rebirth of Bipolar CPUs

Exponential was founded in June 1993 to develop the world's fastest PC processor. While this is an oft-stated goal, the founders had a unique idea to revive the dormant technique of bipolar design. Cofounder George Taylor, now chief technical officer at Exponential, became familiar with the benefits and drawbacks of bipolar

design while working on the ECL-based R6000 at MIPS. Cofounder Jim Blomgren left Sun after it cancelled the ECL SPARC processor he had been working on. Although these ECL processors weren't successful, the founders realized that emerging technologies could be used to address bipolar's problems.

Bipolar design can work with any CPU instruction set, of course, but given the lack of a high-performance PowerPC processor and an offer of funding from Apple, Exponential set out to build a fast PowerPC chip. With \$17 million in venture funding, the company has grown to 60 employees, including seasoned executives such as President Rick Shriner, late of Apple, and VP Marketing Rick Bergman, formerly of Texas Instrument's CPU group. Chairman of the Board Gordon Campbell, founder of Chips & Technologies, has provided guidance and seed funding since the company's inception.

Exponential will be a fabless CPU vendor. The company has contracted with a major semiconductor foundry to build the x704 but would not reveal the identity of this fab partner; sources indicate the fab may be Hitachi, which

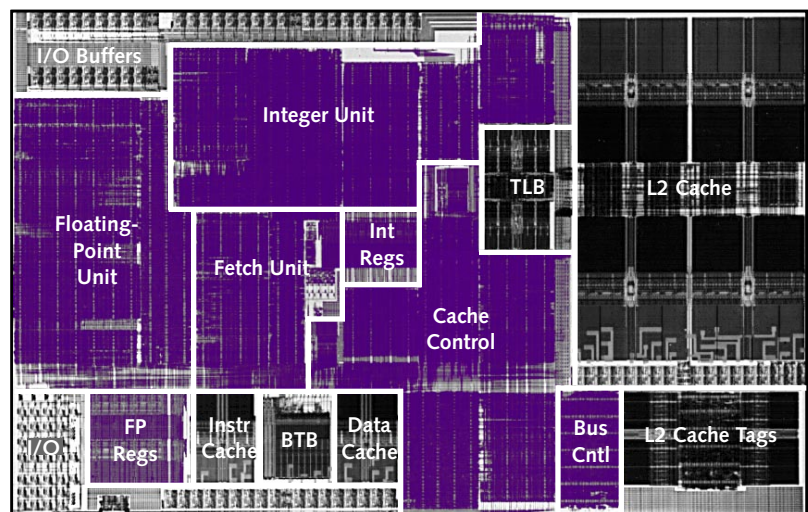


Figure 1. Exponential's x704 combines fast bipolar logic (purple) with dense CMOS memory (black) on a single die. In a 0.5-micron process, the die measures 150 mm².

builds bipolar chips for its mainframes. The startup has obtained a PowerPC license from IBM, shielding it from legal challenges. Given that IBM and Motorola have no PowerPC products that compete with the x704's performance, they are reluctantly willing to welcome the new processor vendor in hopes of expanding the overall market for PowerPC chips.

Traditional Bipolar Chips Are Fast, Hot

Integrated-circuit designers have long known that bipolar logic is faster than CMOS. TTL, ECL, and other bipolar techniques have been used for a variety of small devices, particularly when speed is of the essence. Bipolar circuits are based on junction transistors, which have a higher gain than the field-effect transistors (FETs) used in CMOS circuits.

For example, once the threshold voltage is exceeded, the transistors in the x704 double their output current for every 25-mV increase in the input voltage, allowing a gate to switch with a voltage swing of just 0.25 V. A typical 3.3-V CMOS circuit, in contrast, requires a swing of about 1.5 V to switch states. Moving the voltage across a long metal trace by 0.25 V is inherently faster than moving it by 1.5 V.

Bipolar logic works fine for relatively simple devices but has two drawbacks that affect complex circuits such as microprocessors. First, a bipolar gate draws power continuously; a CMOS gate draws power only when it switches states, and even then it requires a modest amount of current. Thus, the power consumption of a complex bipolar device is much higher than for a comparable CMOS design. Second, a typical bipolar chip contains far fewer transistors than a comparable CMOS chip, in part because the bipolar gates are often less dense, and also because thermal issues often limit the bipolar transistor budget.

Many microprocessor vendors staffed at least one bipolar processor design in the late 1980s. Of these projects, only the R6000 ever reached the market, and that chip had few takers. The CPU designers who explored bipolar discovered that, while a bipolar processor can run faster than a CMOS processor, the CMOS chip gains performance from cramming more memory and logic onto the same size die. These designers found bipolar processor designs to be hotter and more expensive than CMOS ones while delivering little, if any, performance advantage.

For these reasons, most industry watchers, including ourselves, have declared bipolar microprocessors to be dead. MicroUnity's ill-fated foray into bipolar design (see [091402.PDF](#)) did little to reverse this opinion, but Exponential is determined to prove us wrong.

Dealing with High Power

Instead of trying to reduce the power required by its bipolar microprocessor, Exponential simply chose to deal with it. The two problems with an 85-W chip are getting the current in and getting the heat out. With traditional wire bonding, power is delivered only to the edges of the die. A high-power

design would require a huge power grid to move this current to the circuits where it is needed.

Instead, the x704 uses flip-chip bonding, attaching the die directly to the package through small solder bonds scattered across the surface of the die. In this way, power is delivered consistently to all circuits as needed through a minimal power routing network.

The flip-chip attachment also helps get the heat out. Because the die is mounted face down on the top of the package, the back of the die is stuck directly to a heat spreader, providing a low-resistance thermal path. The heat spreader, which is built into the ceramic BGA package, is then attached to a large heat sink that measures 7.0 cm on a side and 2.8 cm high. With this heat sink, an airflow of 23 ft³/minute is required to cool the x704. In contrast, a typical PC fan produces an airflow of 5–10 ft³/minute.

A fan that provides the requisite airflow costs less than \$10, about twice the cost of a standard fan. This fan should be mounted on or next to the processor daughtercard to ensure that all of the air reaches the heat sink. The large heat sink adds another \$8 to the system cost. Neither of these costs is prohibitive, particularly for the high-end systems that will typically use the Exponential chip.

Bipolar logic has completely different power characteristics than CMOS. The power dissipation of the x704 does not scale with clock speed, except for a minor amount of power from the CMOS portions. Thus, the power cannot be reduced by slowing the clock. Similarly, the gated-clock techniques used to reduce power in modern CMOS processors are ineffective for bipolar logic. Furthermore, there is little difference between the typical and maximum power dissipation of the x704, since all the bipolar gates consume power whether the chip is active or not.

The only lever Exponential's designers could use to adjust power dissipation was the design of the transistors themselves. Junction transistors can be designed in a variety of sizes; the bigger the transistor, the more current it can move, and the faster it will switch. For critical speed paths, the designers increased the transistor sizes to improve the clock speed. This trick could be used only so often, however, before the chip reached its power-dissipation limit.

Building Dense Bipolar Circuits

Exponential has done an effective job of solving the density problem. The most obvious technique is to mix ECL and CMOS circuits. Instead of using just a few bipolar transistors, as in a traditional BiCMOS chip such as Pentium, the x704 contains large blocks of pure bipolar logic along with chunks of dense CMOS memory, primarily for the caches. As Figure 1 shows, about half of the x704 die is bipolar logic, with the remainder implemented in CMOS. This combination takes advantage of the high density of CMOS memory cells along with the speed of ECL.

The bipolar regions themselves are also fairly dense. Exponential developed its own library of compact bipolar

devices to help reduce the size of its processor. The x704 also uses a manufacturing process optimized for bipolar transistors, including extra steps for trench isolation (a technique that reduces the die area of bipolar transistors). These extra steps enable bipolar transistors that are about one-fifth the size of bipolar devices built in a typical BiCMOS process. But the extra steps increase wafer processing cost by about a third over pure CMOS, according to the company, and also slightly decrease the chip yield.

Compared with CMOS, ECL also allows more complex functions to be combined into a single gate. For example, a single ECL gate can perform a three-input XOR operation or an 8-to-1 multiplexer with OR inputs. As a result, Exponential claims its ECL gates are roughly the same size as equivalent-function CMOS circuits built in the same process.

Conservative Process Lowers Cost

The x704 is built in a 0.5-micron process, which is to say the CMOS transistors have a drawn gate length of about 0.5 microns; the bipolar transistors have a different geometry. The process includes five metal layers and a local interconnect layer. Even in this fairly conservative process, the x704 measures only 150 mm², about the same size as the original 0.5-micron Pentium or 0.35-micron PowerPC 604e. The chip contains 2.7 million transistors; 2.0 million are in the CMOS memory cells, and 235,000 are bipolar transistors in the memory areas (e.g., in sense amps). The processor core has just 465,000 transistors, all bipolar.

Although Exponential claims its bipolar process is as dense as comparable CMOS processes, the x704 has 6,200 transistors per mm² in its CPU core, about half the density of 0.5-micron CMOS processors such as the 604 and HP's PA-8000. As another example, the PowerPC 601, a similar three-way superscalar processor with 32K of on-chip cache, measures 121 mm² in a 0.7-micron CMOS process; the x704 is bigger despite a more advanced process.

Taking into account the cost of the extra bipolar layers, the 0.5-micron bipolar process should have about the same wafer cost as a 0.35-micron CMOS process. The MDR Cost Model projects the manufacturing cost of the x704 to be about \$90, 50% more than a 0.35-micron PowerPC 604e. According to our model, the Exponential chip costs much less to build than Intel's Pentium Pro, Digital's 21164, the PA-8000, and other high-performance processors.

Our model does not include the cost premium of an external foundry, which Exponential, unlike the aforementioned processor vendors, must pay. This premium could be significant, particularly since Exponential is tied to a single foundry for its unusual IC process.

Two-Level Cache on Chip

Like Digital, Exponential has chosen to optimize performance by maximizing clock speed instead of complexity. Clock speed is the most effective way to increase core CPU performance, as Digital's Alpha processors have demon-

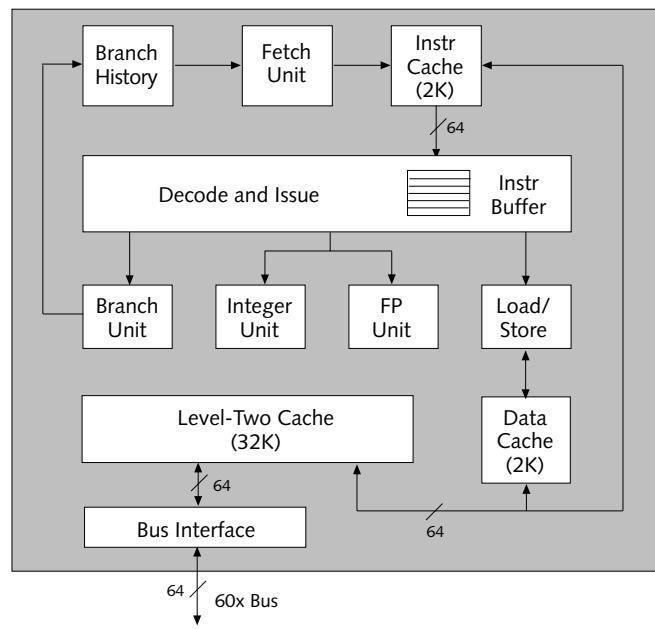


Figure 2. The x704 combines two levels of on-chip cache with a relatively simple three-way superscalar CPU.

strated again and again. Adding complexity reduces the clock speed, typically losing more performance than is gained. The relatively simple x704 design, for example, keeps the critical path length to just 12 gates, enabling the high clock speed. Complexity also increases the cost of manufacturing a chip: as noted, the x704 has a lower manufacturing cost than that of other products in its performance class.

The x704 microarchitecture, shown in Figure 2, is not unlike that of Digital's 21164. The chip has tiny single-cycle primary caches and a larger L2 cache, all on the processor die. The CPU itself issues instructions strictly in order and has no register renaming. The x704 executes up to three instructions per cycle, even on integer code.

The primary instruction and data caches are direct-mapped and only 2K each. Working with a clock period of roughly 2 ns, Exponential ran into the same problem Digital did with the 21164: it was impossible to access a larger cache in such a short time. The 21164 has 8K primary caches but uses a 0.35-micron process to reach 500 MHz; with just 0.5-micron CMOS for the caches, Exponential settled for even smaller caches.

The miss rate of these tiny caches is, of course, terribly high for many applications. Further traffic is generated by the write-through policy of the data cache, although the x704 employs a store queue with merging to reduce this impact. To back up the primary caches, the x704 includes a 32K unified second-level cache on the chip. To improve the hit rate, this cache is eight-way set-associative. The L2 cache has a two-cycle access time but is implemented as two interleaved banks; if accesses are sent to alternate banks, the cache can begin a new access on each cycle.

On each access to the primary data cache, the address is

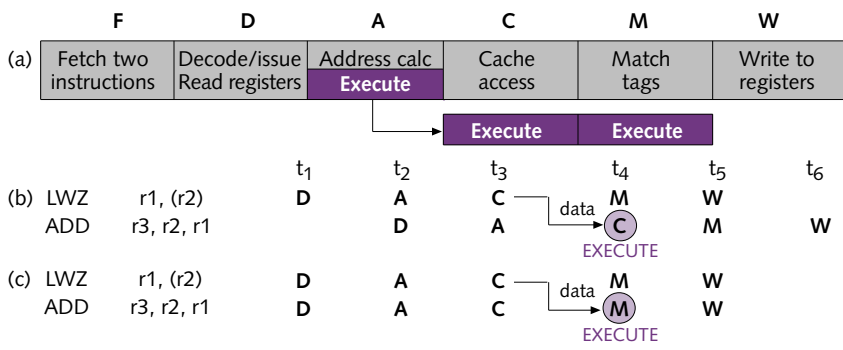


Figure 3. The x704 pipeline has six stages, including an unusual sliding execute stage.

also sent to the L2 cache, where a speculative access is started if the desired bank is available and no higher-priority activity (such as a previous cache refill) needs that bank. If the original access misses in the L1 cache but the speculative access hits in the L2 cache, the miss penalty is three cycles (the two-cycle latency plus one cycle to move the data across the chip). The miss penalty is four cycles if the speculative access cannot be launched. The speculative data access succeeds about half the time, putting the average L1 miss penalty at 3.5 cycles, assuming an L2 hit.

Similarly, the instruction-cache miss penalty is either four or five cycles, depending on the ability to launch a speculative access. Speculative instruction accesses are the lowest priority for the L2 cache, but they still succeed about 30% of the time. The extra cycle on instruction-cache misses allows the incoming instructions to be checked for branches; if a branch is found, the target is calculated and stored in the instruction cache's branch-prediction area.

Superscalar Instruction Issue

Instructions are read from the instruction cache two at a time, under the control of the fetch unit. For each doubleword in the instruction cache (256 entries total), the chip stores two bits of branch history and a predicted target. Using this mechanism, there is no penalty for a correctly predicted taken branch.

This simple prediction method is compact but has a poor prediction rate compared with that of other high-performance processors, due to the small number of entries. The effectiveness is further diluted because many cache doublewords will not contain a branch. Exponential claims a 75% branch-prediction accuracy for SPECint95.

Instructions are fed into a six-entry instruction buffer, where two are decoded per cycle. The third instruction is partially decoded to see if it is a branch. Instructions are issued in order, up to three in a single cycle: one to the load/store unit, one to the arithmetic (integer/FP) unit, and one to the branch unit. Issuing three instructions per cycle is rare, however, as this can occur only if the last instruction in the group is a PC-relative branch. Thus, the rest of the chip is optimized for a sustained rate of two instructions per cycle.

There are few issue restrictions compared with other superscalar processors. Only one instruction can be issued to a function unit on each cycle, and instructions that write to the same register cannot be paired. Instructions need not be in any particular order to be paired. There is only one unusual restriction: a load or store instruction that modifies its base address register (e.g., a postincrement load) is issued to both the load/store and arithmetic units, due to a limited number of write ports on the register file.

Sliding Execute Stage

An unusual feature of the x704 is its sliding execute stage. Figure 3a shows the typical execution pipeline, which has six stages designated F, D, A, C, M, and W. Normally, integer arithmetic and address calculations occur in the A stage, and the data cache is accessed in C. On a cache read, data is available at the end of C and can be used on the next clock cycle, but the tags are not checked until the M stage. If the tags indicate a miss, the incorrect data can be backed out before the register file is written in W.

A number of other processors implement similar pipelines, and the downside is the one-cycle stall that occurs when a load instruction is followed immediately by an instruction that uses the data being loaded. This load-use penalty can also prevent the load instruction from being paired with the use instruction in a superscalar design.

Exponential solves this problem by allowing the execute operation to move down the pipeline into the C or even the M stage. Figure 3b shows a load (LWZ) instruction followed by an ADD instruction that uses the loaded data. In this example, the ADD executes in the C stage, right after the load data is returned from the cache. Note that there is no load-use penalty in this example. In fact, as Figure 3c shows, the load and use instructions can be issued together with no penalty. In this case, the execute stage moves into M. Once the execute stage is in M, subsequent load and use instructions can continue to be paired without any penalty.

If the execute stage has moved to C or M, it will attempt to return to its original position on subsequent cycles. It can move backwards, however, only during a cycle in which there is no instruction issued to the arithmetic unit. Since this unit is rarely unused, the execute stage tends to slide out to C or M and stay there until there is a mispredicted branch.

The drawback of sliding out the execute stage is that it delays the availability of results. For most calculations, this is not a problem, since the execute stage remains in its delayed position for as long as necessary. The problem is mispredicted branches: the later the execute stage, the longer it takes before the processor realizes the branch is mispredicted.

If the execute stage is in A, the mispredicted branch penalty is three cycles. If it has been pushed all the way to M,

the penalty is five cycles. This penalty can be reduced to a minimum of two cycles if the PowerPC condition-code instruction is separated from the branch instruction.

The sliding execute stage thus provides a dynamic tradeoff. In code with few instructions between branches, the execute stage tends to stay in A, minimizing the misprediction penalty. With lots of instructions between branches, the execute stage is likely to end up in M, but in this case the impact of a longer misprediction penalty is lessened by the reduction in the number of branches. The upside is the elimination of the load-use penalty even when the load and use instructions are paired. Exponential says the overall impact of the sliding execute stage is a 3–5% performance improvement.

Compatible with 60x Bus

The x704 is designed to be compatible with the 60x bus used by all other PowerPC chips, including support for burst-mode transactions and multiprocessor configurations. Thus, it should interface with commercially available system-logic chip sets from Motorola and others as well as with proprietary system logic developed by Apple and IBM. Most of Exponential's development testing has been done by plugging x704 daughtercards into stock Macintosh systems.

The x704 is not a simple pin-compatible replacement, however. First, the chip uses a 356-pin ceramic BGA package instead of the 604's 225-pin BGA. The extra pins are mainly for added power and ground current. Second, the bipolar logic requires two unusual voltage inputs: the primary supply is 3.6 V, while a secondary 2.1-V input is used for the termination voltage.

Most current PowerPC systems place the processor on a daughtercard. For these systems, no modifications are required for the motherboard. The x704 daughtercard contains the CPU along with the voltage regulators and power FETs required to generate the appropriate voltages from the system's 5-V supply. The x704 requires no special clocking: it has an on-chip PLL that generates the CPU clock as any integer multiple (4× to 17×) of the bus clock.

Other modifications may need to be made to the system itself. The power supply must have about 60 W of extra capacity. Many systems have plenty of margin in this area, and upgrading to a higher-capacity commodity supply is easy and relatively inexpensive; some Macintoshes, unfortunately, do not use commodity power supplies. An extra fan is usually required to ensure adequate airflow across the toasty processor, and some internal ducts or baffles may be needed to properly channel the air.

Interestingly, the x704 has an on-chip thermal sensor designed to warn the system if the chip overheats (assuming the system software has been modified to handle this task) and to shut down the processor if the maximum temperature is exceeded. This sensor should avert damage to the chip in the event of a fan failure or a poor system design. Unlike a

	Single Precision		Double Precision	
	Throughput	Latency	Throughput	Latency
Int Multiply	3–6 cycles	3–6 cycles	n/a	n/a
Int Divide	37 cycles	37 cycles	n/a	n/a
FP Add, Sub	1 cycle	4 cycles	1 cycle	4 cycles
FP Multiply	1 cycle	4 cycles	2 cycles	5 cycles
FP Mul-Add	1 cycle	4 cycles	2 cycles	5 cycles
FP Divide	20–66 cyc	21–67 cyc	34–140 cyc	35–141 cyc

Table 1. For most operations, the FPU is fully pipelined with a four-cycle latency; divides and double-precision multiplies take longer. n/a indicates not applicable.

CMOS processor, the x704 can't simply run at a reduced clock speed in these situations, as this would not reduce its heat dissipation; instead, it must power down.

In summary, designing the x704 into an existing system should be relatively straightforward. The component cost of the voltage regulators, bigger power supply, and extra fan should be less than \$40. New quieter fans should avoid any significant increase in user-perceived noise. Offering the x704 as a field upgrade would be difficult, however, unless the system had been designed for this from the start.

Bandwidth May Limit Performance

For at least some applications, Exponential's coupling of a 533-MHz CPU with the 60x bus will be like strapping a jet engine to a Piper Cub. The x704 has roughly twice the core performance of a 200-MHz 604e but half as much on-chip cache. Even considering the greater associativity of the x704's cache, the faster chip will generate up to twice as many cache misses as the 604e on the same workload. Unfortunately, for many applications today, a fast 604e already comes close to maxing out the bandwidth of the 60x bus.

Some gains can be seen by using a higher-speed system design. The x704 can drive its external bus at up to 100 MHz, connecting directly to a large external synchronous cache. An external chip set could provide cache-control logic while connecting to a 66-MHz 60x bus for main-memory and I/O traffic. Most current 60x systems, however, connect the CPU directly to a 66-MHz (or slower) 60x bus.

For relatively small benchmarks, including SPECint95, the x704 should perform well even in a typical Macintosh design. For larger applications, however, performance is likely to suffer without a large external cache running at more than 66 MHz.

For small floating-point applications, performance should be comparable to that of other high-end PowerPC chips. As Table 1 shows, the x704 completes all single-precision floating-point operations, except division, in four clock cycles with single-cycle throughput. Double-precision multiplication takes an extra cycle of throughput and latency. These figures are similar to the 604's but not as good as those of other high-end RISC processors. The bandwidth limitations of the 60x bus will be particularly acute for most FP applications, which often have large data sets. Exponential is focused mainly on integer-only PC software.

Price & Availability

The x704 is expected to be offered in speeds of 466, 500, and 533 MHz in a 356-lead BGA package. Samples are due in 1Q97, with volume production in 2Q97. Exponential has not yet announced list pricing but indicated it may be around \$1,000. For more information, contact Exponential (San Jose, Calif.) at 408.441.6050 or on the Web at www.exp.com.

World's Fastest PC Processor

The company received first silicon of the x704 last May and is well along with application testing and verification. At the Forum, Exponential's Taylor said volume production and system announcements would occur sometime in 2Q97. Given the progress to date, this forecast seems quite reasonable.

Taylor projects the 533-MHz x704 will deliver 11–13 SPECint95 and 10 SPECfp95 (base). These estimates are based on comparisons to the PowerPC 604; the company has not measured SPEC performance. In fact, since the startup's targets are the Macintosh and Windows NT markets, it may never measure SPEC performance.

Assuming the company's estimates are accurate, the x704's integer performance would put it ahead of any microprocessor shipping today, with the possible exception of the 21164. The FP score is well behind that of workstation RISC processors but is comparable to the best available from Intel or from other PowerPC processors. Exponential's stunted cache design may limit its chip's performance on sizable PC applications, which often stress the cache more than SPECint95.

By the time the x704 is available in 2Q97, RISC processors from HP, MIPS, and Sun are expected to be in the same SPECint95 range as Exponential's device. The x704 should still be ahead of Intel's Klamath and anything in the PowerPC line. By the end of 1997, however, the so-called G3 PowerPC (see [101103.PDF](#)) and Intel's Deschutes processor should match the x704's performance.

Exponential plans to offer the x704 at speeds of 466, 500, and 533 MHz. Bipolar circuits have a narrow speed yield, so the range of product offerings is not broad. The company has not yet announced pricing for the parts but believes it will be close to \$1,000, at least initially. Since the x704's performance will be unmatched in the PC space when it is released, the company will be able to justify this premium price.

At this price, however, the x704 will not offer a significant price/performance advantage over Intel processors, doing little to make PowerPC more attractive to fence sitters.

A small company with a need to quickly reach a positive cash flow, Exponential appears intent on milking existing PowerPC system vendors rather than taking an aggressive pricing stance that might attract new players.

Once PC processors like Deschutes and the G3 near the x704's performance, the company will have to cut its price significantly. Fortunately, the modest manufacturing cost of the 0.5-micron device should enable Exponential to make a profit even at a price of a few hundred dollars.

The x704's position in the NT workstation market is more difficult, mainly because of Alpha. Digital's 500-MHz 21164 already offers integer performance comparable to the x704's estimates, and by next spring the Alpha chip may be even faster. Digital's 21164PC (see [1014MSB.PDF](#)), slated to be out at the same time as Exponential's chip, will offer similar performance in the same low-cost system configurations at which the x704 is aimed. Thus, Exponential will not have a clear performance lead for NT.

To gain traction in the NT market, the startup should price more aggressively and promote the PowerPC common hardware platform, with its Macintosh and Windows NT compatibility. Unfortunately, few software vendors have ported NT applications to PowerPC, and even fewer system vendors are offering NT on PowerPC. Thus, we suspect Exponential will instead keep its pricing focus on the Macintosh market and view any NT sales as incidental.

Roadmap to Higher Performance

As it nears the completion of its first product, Exponential has plans to push performance even further in future generations. Next up is a processor code-named X2, expected to ship by mid-1998, about a year after the first chip. While using the same 0.5-micron bipolar process, the X2 will offer higher clock speeds, larger on-chip caches, and a backside L3 cache bus to break the bandwidth bottleneck. These changes will increase the die size and, more important, the power dissipation; without a process change, the higher clock speeds will be achieved by trading power for speed in the bipolar logic.

A process shrink is slated for mid-1999. The new process should allow even higher clock speeds and larger on-chip caches while easing power dissipation. This chip, the X3, will include a new system bus, probably compatible with the G3's, but no significant microarchitecture enhancements.

A key issue for Exponential in the long run is the ability of its bipolar process to scale as IC process geometries decrease. Bipolar gates do not operate well at 2.5 V or below, and most 0.25-micron processes are being designed for these low voltages (see [101203.PDF](#)). In addition, at these voltages, the signal swing of a CMOS gate is roughly the same as that of a



MICHAEL MUSTACCHI

Exponential cofounder George Taylor unveils a 533-MHz PowerPC chip at the Microprocessor Forum.

bipolar gate, eliminating much of the speed advantage of bipolar. Despite these trends, Exponential believes its bipolar designs will have a long-term advantage.

This advantage is critical to Exponential's future success: except for the sliding execute stage, the x704 microarchitecture has no unusual advantages. The chip gains its performance advantage from the high clock speed, which in turn is due to the bipolar process. The startup needs this process advantage and will be in trouble if its foundry chooses not to develop new bipolar processes or not to build Exponential's chips for any reason.

Exploiting an Opening

Exponential's bipolar technology takes advantage of a neglected tradeoff in system design: heat for performance. System designers willing to take the time and spend a few extra dollars to support an 85-W processor can be rewarded with performance otherwise unobtainable in the PC space. The

company has also cleverly targeted an opening in the high end of the PowerPC product line, one that Motorola and IBM have so far been unable to fill.

In the short term, Exponential's challenge is to convince vendors to redesign their systems not just for cooling but to add the bus bandwidth required to make effective use of its powerful CPU. We expect Apple and others will make this effort in order to place high-performance, high-margin systems at the top of their product lines.

In the long term, the startup's position is more tenuous. Both Intel, with Merced, and the IBM/Motorola partnership, with G3 and G4, are making significant efforts to upgrade their performance over the next two years. Exponential will have to work hard to maintain a performance edge over these chips, particularly if its bipolar process advantage begins to wane. But with its bold initial effort, Exponential should prosper over the next couple of years while girding itself for the battles yet to come. 