

# making a case for Efficient Super

WU-CHUN FENG,  
LOS ALAMOS NATIONAL LABORATORY

A supercomputer evokes images of “big iron” and speed; it is the Formula 1 racecar of computing. As we venture forth into the new millennium, however, I argue that efficiency, reliability, and availability will become the dominant issues by the end of this decade, not only for supercomputing, but also for computing in general.

Over the past few decades, the supercomputing industry has focused on and continues to focus on performance in terms of speed and horsepower, as evidenced by the annual Gordon Bell Awards for performance at Supercomputing (SC). Such a view is akin to deciding to purchase an automobile based primarily on its top speed and horsepower. Although this narrow view is useful in the context of achieving “performance at any cost,” it is not necessarily the view that one should use to purchase a vehicle. The frugal consumer might

It's time for the computing  
community to use alternative metrics for  
evaluating performance.

# computing

# making a case for Efficient Supercomputing

## FIG 1

Courtesy of Lawrence Livermore National Library.



of new machine rooms, as shown in figure 1, and in some cases, requires the construction of entirely new buildings. The primary reason for this

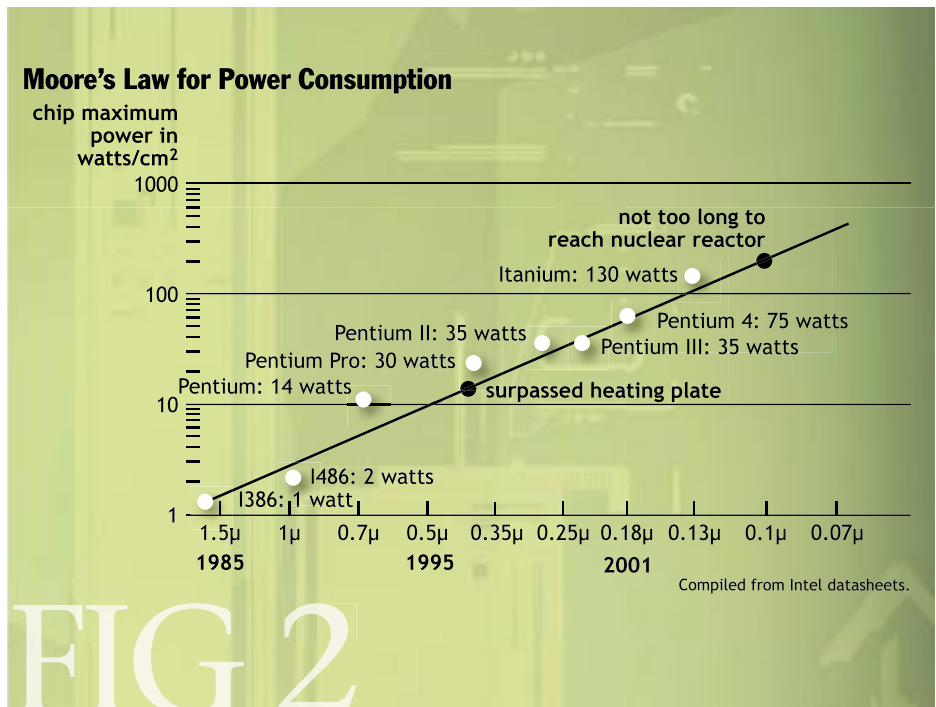
consider fuel efficiency, reliability, and acquisition cost. Translation: Buy a Honda Civic, not a Formula 1 racecar. The outdoor adventurer would likely consider off-road prowess (or off-road efficiency). Translation: Buy a Ford Explorer sport-utility vehicle, not a Formula 1 racecar. Correspondingly, I believe that the supercomputing (or more generally, computing) community ought to have alternative metrics to evaluate supercomputers—specifically metrics that relate to efficiency, reliability, and availability, such as the total cost of ownership (TCO), performance/power ratio, performance/space ratio, failure rate, and uptime.

### MOTIVATION

In 1991, a Cray C90 vector supercomputer occupied about 600 square feet (sf) and required 500 kilowatts (kW) of power. The ASCI Q supercomputer at Los Alamos National Laboratory will ultimately occupy more than 21,000 sf and require 3,000 kW. Although the performance between these two systems has increased by nearly a factor of 2,000, the performance per watt has increased only 300-fold, and the performance per square foot has increased by a paltry factor of 65. This latter number implies that supercomputers are making less efficient use of the space that they occupy, which often results in the design and construction

less efficient use of space is the exponentially increasing power requirements of compute nodes, a phenomenon I refer to as “Moore’s law for power consumption” (see figure 2)—that is, the power consumption of compute nodes doubles every 18 months. This is a corollary to Moore’s law, which states that the number of transistors per square inch on a processor doubles every 18 months [1]. When nodes consume and dissipate more power, they must be spaced out and aggressively cooled.

Without the exotic housing facilities in figure 1, traditional (inefficient) supercomputers would be so unreliable (due to overheating) that they would never be available for use by the application scientist. In fact, unpublished empirical data from two leading vendors corroborates that the failure rate of a compute node doubles with



## FIG 2



every 10-degree C (18-degree F) increase in temperature, as per Arrhenius' equation when applied to microelectronics; and temperature is proportional to power consumption.

We can then extend this argument to the more general computing community. For example, for e-businesses such as Amazon.com that use multiple compute systems to process online orders, the cost of downtime resulting from the unreliability and unavailability of computer systems can be astronomical, as shown in table 1—millions of dollars per hour for brokerages and credit card companies and hundreds of thousands of dollars per hour for online retailers and services. This downtime cost has two components: lost revenue (e.g., the end user “clicking over” to the competitor’s Web site) and additional hours of labor spent fixing the computer systems.

**Table 1. Estimated Costs of an Hour of Server Downtime for Business Services**

Service	Cost of One Hour of Downtime
Brokerage Operations	\$6,450,000
Credit Card Authorization	\$2,600,000
eBay	\$225,000
Amazon.com	\$180,000
Package Shipping Services	\$150,000
Home Shopping Channel	\$113,000
Catalog Sales Center	\$90,000

Source: Contingency Planning Research.

Clearly, downtime should be a component in the total cost of ownership (TCO) of a computer system, whether the system is a Web-server farm or a supercomputer. But what other components make up TCO? More generally than even downtime, TCO consists of two parts: (1) cost of acquisition and (2) cost of operation. The former is a one-time cost that can be defined as all the costs incurred in acquiring a computer system—for example, procurement, negotiation, and purchase—and, thus, is relatively straightforward to quantify [2]. The latter, however, is a recurring cost that consists of multiple components, including costs related to system integration and administration, power and cooling, downtime, and space. Although the costs related to power and cooling and space are easily quantifiable, the other operational costs—that is, system integration and administration and downtime—tend to be highly institution-specific and full of hidden costs [3]. As a result, I conclude that TCO *cannot* be easily quantified. I instead focus on quantifying metrics that are related to TCO such as efficiency, reliability, and availability. Specifically, we propose

the following metrics: performance/power ratio, performance/space ratio (or compute density), failure rate, and uptime.

#### EFFICIENT SUPERCOMPUTING

Green Destiny, as shown in figure 3, is the name of our 240-processor supercomputer that fits in a telephone booth and sips less than 5.2 kW of power at full load (and only 3.2 kW when running diskless and computationally idle). It provides affordable, general-purpose computing to our application scientists while sitting in an 85- to 90-degree F dusty warehouse at 7,400 feet above sea level. More importantly, it provides reliable computing cycles without any special facilities—that is, no air conditioning, no humidification control, no air filtration, and no ventilation—and without any downtime. (In contrast, a more traditional, high-end 240-processor supercomputer such as a Beowulf cluster [4] generally requires a specially cooled machine room to operate reliably, as such a supercomputer easily consumes as much as 36.0 kW of power and cooling, roughly seven times more than Green Destiny.)

Green Destiny takes a novel and revolutionary approach to supercomputing, one that ultimately redefines performance to encompass metrics that are of more relevance to end users: efficiency, reliability, and availability. As such, Green Destiny is arguably the world’s most efficient supercomputer as it provides a completely integrated solution that is orders of magnitude

FIG 3

Courtesy of Los Alamos National Laboratory.



# making a case for Efficient Supercomputing

superior to any other solution based on efficiency, reliability, availability, versatility, management, self-monitoring and measurement, and ease of use [5,6].

## THE MAGIC BEHIND GREEN DESTINY

To achieve such efficiency, reliability, and availability, we designed an architecture around which we could appropriately stitch together the modified building blocks of Green Destiny. These building blocks include a Transmeta-powered RLX ServerBlade as the compute node and World Wide Packets' Lightning Edge network switches configured in a one-level tree topology for efficient communication, as shown in figure 4.

By selecting a Transmeta processor as a compute engine, Green Destiny takes a predominantly hardware-based approach to power-aware supercomputing. A Transmeta processor eliminates about 75 percent of the transistors used in a traditional RISC architecture and implements the lost (but inefficient) hardware functionality in its code-morphing software (CMS), a software layer that sits directly on the Transmeta hardware. This approach results in a processor that runs cooler than other processors, as illustrated by figure 5, which shows the thermal images of a conventional, low-power, mobile processor and a Transmeta processor. The operating

temperatures of the processors differ by 57.3 degrees C (or 135.1 degrees F) when running a software-based DVD player. This means that based on the corroborated Arrhenius equation, the conventional, low-power, mobile processor (without any active cooling) is 32 times more likely to fail than the Transmeta processor (without any active cooling).

Although the Transmeta processor is significantly more reliable than a conventional mobile processor, its Achilles' heel is its floating-point performance. Consequently, we modified the CMS to create a "high-performance CMS" that improves floating-point performance by nearly 50 percent and ultimately matches the performance of the conventional mobile processor on a clock-cycle-by-clock-cycle basis.

On the network side, Green Destiny runs a software configuration for the Lightning Edge switches where features such as auto-negotiation are simply turned off, since all link speeds are known. This reduces power consumption down to a few watts per port.

## APPLICATIONS FOR GREEN DESTINY

Initially, we turned to the theoretical astrophysics community for a scientific application to run on Green Destiny: an n-body simulation containing a few hundred-thousand galaxies [4], as shown in figure 6. This application was followed up with a smoothed particle hydrodynamic simulation of a three-dimensional supernova core-collapse [5]. Since then, we have also run applications in the fields of large-scale molecular dynamics and bioinformatics. For the latter, we developed our own parallel BLAST code for sequence matching called mpiBLAST [7], a code that demonstrates super-linear speedup.

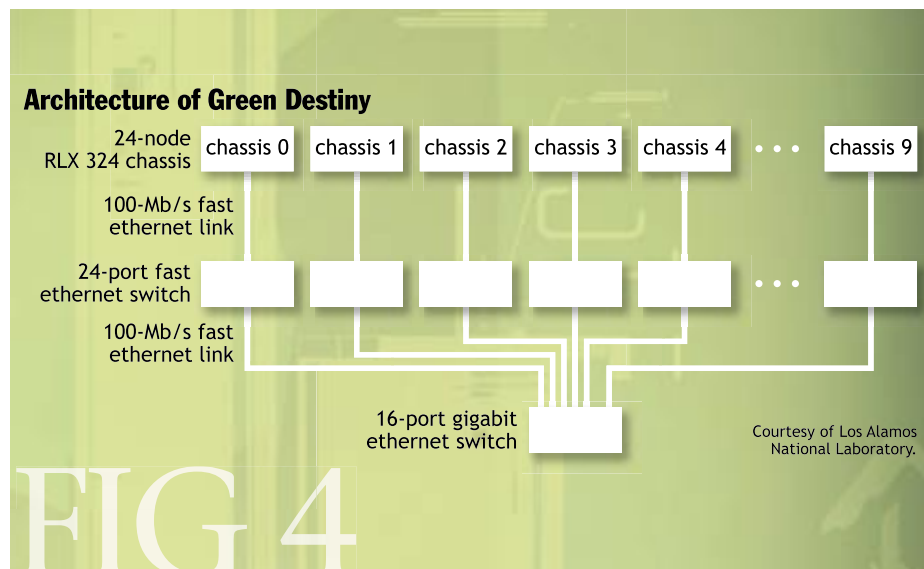
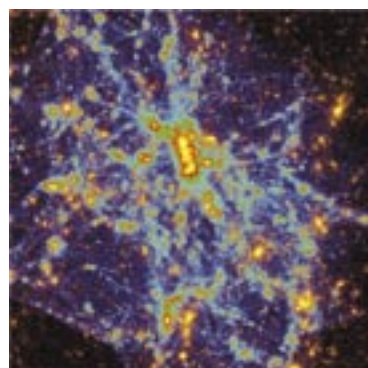


FIG 5



Courtesy of Transmeta.

FIG 6



Courtesy of Los Alamos National Laboratory.

Figure 6 shows an intermediate stage of a gravitational n-body simulation of a galaxy formation with 10 million particles. The overall simulation of 1,000 timesteps with more than  $10^{15}$  floating-point operations completed in less than one day on Green Destiny. The region shown in figure 6 represents about 150 million light-years across.

The most time-consuming part of this application is computing components of the accelerations of particles [8], in particular, evaluating  $r^{-3/2}$  where  $r$  is the separation between particles. Because of the importance of this calculation to general n-body codes, we evaluate the uniprocessor performance of commodity processors using two different implementations of a reciprocal square-root function—(1) *sqrt* function from a math library and (2) Karp's implementation of square root [8]—as part of a gravitational microkernel benchmark. To simulate the calculation in the context of an n-body simulation (and, coincidentally, enhance the confidence interval of our floating-point evaluation), our gravitational microkernel benchmark loops 100 times over the reciprocal square-root calculation.

Table 2 shows the Mflops ratings for six commodity processors over the two different implementations of

the gravitational microkernel benchmark, where Mflops stands for mega ( $10^6$ ) floating-point operations per second. Considering that the Transmeta processors are software-hardware hybrids and the other processors are all-hardware designs, the Transmetas with our high-performance CMS run remarkably well.

**Table 2. Mflops Ratings on a Gravitational Microkernel Benchmark**

Processor	Math sqrt (libm)	Karp sqrt
500-MHz Intel Pentium III	87.6	137.5
533-MHz Compaq Alpha EV56	76.2	178.5
667-MHz Transmeta TM5600	128.7	297.5
933-MHz Transmeta TM5800	189.5	373.2
375-MHz IBM Power3	298.5	379.1
1,200-MHz AMD Athlon MP	350.7	452.5

Note: Larger Mflops ratings are better.

The computational efficiency of each processor with respect to power, hereafter referred to as *power efficiency of the processor*, is shown in table 3. Given that the uniprocessor performance of the Transmeta is comparable to traditional power-hungry processors, table 3 provides motivation for the computing industry to expand its horizons to address performance from a mixture of at least two perspectives: speed and power efficiency.

**Table 3. Power Efficiency on a Gravitational Microkernel Benchmark (Larger performance/power ratios are better)**

Processor	Performance/Power Ratio	
	Math sqrt (libm)	Karp sqrt
500-MHz Intel Pentium III	5.1	8.0
533-MHz Compaq Alpha EV56	0.85	2.0
667-MHz Transmeta TM5600	17.6	40.8
933-MHz Transmeta TM5800	31.6	62.2
375-MHz IBM Power3	37.3	47.4
1,200-MHz AMD Athlon MP	6.2	8.0

Note: The power consumption that is used for each of these processors is based on the manufacturers' data sheets.

Table 4 provides a historical account on the performance of supercomputing clusters running a standard n-body simulation, starting from a spherical distribution of particles that represents the initial evolution of a cosmological n-body simulation [6]. Somewhat surprisingly, we find that the performance per processor for Green Destiny (667-MHz Transmeta TM5600-based) and Green Destiny+ (933-MHz/1-GHz Transmeta TM5800-based) on this parallel n-body code is substantially better than the SGI Origin 2000 supercomputer and comes

# making a case for Efficient Supercomputing

within 10 percent of matching the performance per processor on ASCI White, a supercomputer that currently ranks in the top 10 of the Top 500 Supercomputer List. (Note: For both Green Destiny and Green Destiny+, we used our high-performance CMS, which improved per-node performance by 50 percent over the standard CMS.)

Though table 4 provides interesting performance numbers for comparative purposes, combining these performance numbers with other known (or measured) quantities such as power consumption and footprint size produces a plethora of even more provocative data points with respect to efficiency: memory density, disk

**Table 4. Historical Performance of n-body Treecode on Clusters and Supercomputers**

Site	Machine	Processor (Proc)	# Procs	Gflops	Mflops/Proc
LLNL	ASCI White	IBM Power3	8,192	2,500	305
LANL	Green Destiny+	Transmeta TM5800	212	58	274
LANL	SGI Origin 2000	MIPS R10000	64	13	203
LANL	Green Destiny	Transmeta TM5600	212	39	184
SC'01	MetaBlade2	Transmeta TM5800	24	3	125
LANL	Avalon	DEC Alpha 21164A	128	16	125
LANL	MetaBlade	Transmeta TM5600	24	2	83
NAS	IBM SP-2(66/W)	IBM SP-2	128	10	78
SNL	ASCI Red	Intel Pentium Pro	6,800	465	68
LANL	Loki	Intel Pentium Pro	16	1	63
SC'96	Loki+Hyglac	Intel Pentium Pro	32	2	63
Caltech	Naegling	Intel Pentium Pro	96	6	63
NRL	TMC CM-5E	Sun SuperSPARC	256	12	47
SNL	ASCI Red	Intel Pentium Pro	4,096	164	40
JPL	Cray T3D	Cray	256	8	31
LANL	TMC CM-5	Sun SPARC2	512	14	27
Caltech	Intel Paragon	Intel iPSC/860	512	14	27
Caltech	Intel Delta	Intel i860	512	10	20

Note: Gflop = giga (10<sup>9</sup>) floating-point operations per second. The Gflop ratings are rounded to the nearest integer Gflop.

**Table 5. Performance and Efficiency Numbers for Clusters and Supercomputers**

Machine	Avalon (1996)	ASCI Red (1996)	ASCI White (2000)	ASCI Q (2002)	Green Des- tiny+(2002) <sup>12</sup>
Performance (Gflops)	18	600	2,500	8,000	58
Memory (GB)	36	585	6,200	12,000	150
Disk (TB)	0.4	2	160	600	5
Area (sf)	120	1,600	9,920	21,000	6
Power (kW)	18	1,200	2,000	3,000	5
Memory Density (MB/sf)	307	374	640	585	25,600
Disk Density (GB/sf)	3	1	17	29	853
Compute Density (Mflops/sf)	150	375	252	381	9,667
Power Efficiency (Mflops/watt)	1.0	0.5	1.2	2.7	11.6

Note: The performance numbers above are based on an actual run of an n-body treecode.

density, compute density (or space efficiency), and power efficiency, as shown in table 5 [9]. The memory density of Green Destiny is more than 40 times better than its closest competing supercomputer; its disk density is 30 times better; its compute density (i.e., performance/space ratio) is 25 times better; and its power efficiency (i.e., performance/power ratio) is roughly 5 to 10 times better than all the other supercomputing platforms.

Note, however, that the comparison in table 5 is a bit of an apples-to-tangerines-to-oranges comparison. The “apple” is Green Destiny, whose purpose is to provide super-efficient and highly reliable supercomputing at the expense of some performance—that is, the Toyota Camry of computing [10]. Based on the data in table 5, Green Destiny clearly makes the most efficient use of space and power (see the Green Destiny colored numbers in table 5). The “oranges” are the ASCI machines with the sole

purpose of achieving performance at any cost—that is, the Formula 1 racecar of computing. Given that the ASCI Q machine leads in every such category (see the ASCI Q colored numbers in table 5), it clearly achieves that purpose. The “tangerine” is Avalon, one of the first Beowulf clusters [4] built with the Linux operating system [11]. Its purpose is to deliver the best price/performance ratio, where price is defined as the cost of acquisition. Of all the supercomputers, Avalon does indeed achieve the best price/performance ratio, just edging out Green Destiny.

To explore yet another intriguing (but still apples-to-oranges) comparison, we look at the LINPACK runs of Green Destiny+ and the Japanese Earth Simulator [13] in table 6. The performance of Green Destiny+ is extrapolated from the measured performance on smaller versions of the machine with the same architecture. The extrapolation is realistic, as the 101-Gflop rating is based on the percentage of peak performance achieved on each of the smaller versions of Green Destiny—that is, 70 percent of peak achieved with LINPACK.

#### FUTURE DIRECTIONS FOR POWER-AWARE SUPERCOMPUTING

Green Destiny represents a primarily *hardware-driven* (or “architecturally driven”) approach to power-aware supercomputing. Such an approach targets a new fabrication technology or hardware redesign with the same functionality but lower energy costs. Two alternative approaches that warrant further investigation are:

- A *software-driven* approach that relies on the hardware to provide an interface that the software can use to control a processor’s frequency and/or voltage levels (thus controlling power consumption as well, because power is directly proportional to the processor’s clock frequency and the square of its operating voltage).
- A *hardware–software* codesign approach that combines

**Table 6. Performance and Efficiency Numbers for Green Destiny+ and the Japanese Earth Simulator**

Machine	U.S. Green Destiny+ [14]	Japanese Earth Simulator
Performance (Gflops)	101	35,860
Memory (GB)	150	10,000
Disk (TB)	5	n/a
Area (sf)	6	70,290 [15]
Power (kW)	5	7,000
Memory Density (MB/sf)	25,600	146
Disk Density (GB/sf)	853	n/a
Compute Density (Mflops/sf)	16,833	510
Power Efficiency (Mflops/watt)	20	5



**We believe that the hardware–software codesign approach holds the most promise.**

both approaches in the hopes of maximizing energy savings while minimizing the impact on overall performance.

At the present time, the architecturally driven approach is the most mature of the three approaches. Hardware designers in embedded computing have been implicitly working in this area for many years, if not decades, to meet the electrical and thermal specifications (or envelopes) of material goods. However, it is still in its infancy with respect to supercomputing and high-performance computing. Primary limitations with this approach include its inflexibility to new technological advances and the unknown requirements from software when designing the hardware.

A software-driven approach must address two control issues: (1) when to direct the hardware to switch between two different levels of power—that is, voltage and/or frequency; and (2) how to rearrange application software code, thus altering the system load and allowing further low-power optimizations to be made. Both of these issues can be quite expensive, particularly without knowledge of the underlying hardware architecture. For example, the decision to switch voltage or frequency levels may



# making a case for Efficient Supercomputing

depend on the hardware overhead involved in performing the switch.

We believe that the hardware-software codesign approach holds the most promise, but it will require significant cooperation between the hardware and software. Power-aware interfaces between the hardware and software will enable the operating-system (OS) programmer to introduce power awareness into traditional OS services. These power-aware OS interfaces must then be made accessible to application programmers so that application-specific information can be transferred to the OS to enable even more effective power management.

## WHY GREEN DESTINY?

Green Destiny provides a completely integrated solution that is orders of magnitude superior to any other solution based on efficiency, reliability, and availability. Specifically, as seen in table 5, its memory density is 40 to 80 times better than traditional supercomputers; its disk density is 30 to 850 times better; its compute density or space efficiency (i.e., performance/space ratio) is 25 to 60 times better; and its power efficiency (i.e., performance/power ratio) is roughly 5 to 10 times better. (Perhaps, an alternative name for Green Destiny could be *Green Density*).

Furthermore, because of its low-power design, Green Destiny has never failed in its lifetime, and

its uptime has effectively been 24 hours a day, 7 days a week, 365 days a year. This means that no time, no effort, and no money were wasted on personnel to diagnose and fix a failure or set of failures; no money was wasted on replacing hardware parts; and Green Destiny was always available for use. This is in direct contrast to our previous supercomputer, a traditional 128-processor cluster that failed on a weekly basis and required as much as a half to full day to diagnose and fix. Even more amazing, Green Destiny manages to achieve all these virtues while operating

in an 85- to 90-degree F dusty warehouse at 7,400 feet above sea level. Although the total cost of ownership is not explicitly addressed in this discussion, it should be clear that the TCO of Green Destiny would be substantially better than any other supercomputing platform.

As noted recently by C. Gordon Bell—sponsor of the Gordon Bell High-Performance Computing Awards at Supercomputing, inventor of the VAX series of minicomputers at Digital Equipment Corporation (DEC), and senior researcher at Microsoft Bay Area Research Center—Green Destiny has stunned the computing industry by “redefining the accessibility and economics of supercomputing to the masses [16].” Further support for Green Destiny comes from J. Craig Venter, founder of Celera Genomics, who stated that he had to spend as much money on his

EnergyGuide labels show the estimated yearly electricity consumption to operate a product along with a scale for comparison with similar products. The comparison scale shows the least and most energy used by comparable models. The EnergyGuide allows consumers to compare the labeled model with other similar models.

Among supercomputers, the 240-processor Green Destiny cluster uses the least amount of energy per hour: 5.2 kilowatt hours (kWh) at load with disks (or 3.2 kWh running diskless but computationally idle).

This label is required to appear on appliances, but not on computers.

Courtesy of Los Alamos National Laboratory.

Alpha supercomputer (\$6 million) as on the appropriate infrastructure to house the supercomputer (\$6 million) in the race to sequence the human genome. As he noted in his interview with GenomeWeb on Oct. 16, 2002 [17], if this is what the bioinformatics revolution was going to cost, then it would be a revolution that would not go very far. This is the primary reason for his interest in the “green machines” that we have developed at Los Alamos National Laboratory.

## CONCLUSION

Green Destiny is merely the first (and hopefully, not last) step in power-aware supercomputing. Its success—particularly in the applications community where specially cooled, machine-room infrastructures are a rarity—stems directly from eschewing Moore’s law with respect to power consumption. Rather than using processors that consume upwards of 100 watts per centimeter (as “prescribed” by Moore’s law for power consumption in figure 2), we based Green Destiny on low-power building blocks—for example, Transmeta processors that consume only six watts per centimeter at load and World Wide Packet switches that consume only a few watts per port. The less power a processor draws, the cooler it will run. The cooler a processor runs, the less likely the overall system will fail (or clock down). By aggressively pursuing a “cool” supercomputer, we ran Green Destiny without any failures in a hostile environment—that is, 85- to 90-degree F in a dusty warehouse at 7,400 feet above sea level with no facilities for cooling, humidification control, or air filtration. In contrast, traditional supercomputers are now so large and use so much power that institutions often construct new machine rooms (and sometimes even new buildings) to house them.

Although I believe that Moore’s law is technically feasible through 2010, and very likely beyond that, its current trajectory is slated to reach one kilowatt per square centimeter by 2010, which is allegedly as much power per square centimeter as the surface of the sun! From a socioeconomic viewpoint, I believe that we must avoid Moore’s law and redirect the performance evaluation of supercomputing systems to metrics other than performance and price/performance. In this discussion, I suggested a few such metrics: total cost of ownership, performance/space ratio, performance/power ratio, reliability, and uptime. A more controversial metric would be “total price/performance ratio” (ToPPer), where total price is defined by TCO. Details about this metric can be found in “The Bladed Beowulf: A Cost-Effective Alternative to Traditional Beowulfs,” an article I coauthored with

Michael Warren and Eric Weigle [5].

By applying these arguments to more traditional data centers such as search-engine farms (e.g., Google), Web-server farms (e.g., Yahoo), and compute-server farms (e.g., IBM’s On-Demand and Hewlett-Packard’s Demand More), even greater economic and efficiency benefits are to be gained. From an efficiency standpoint, the computational density, memory density, disk density, and power efficiency of Green Destiny is at least an order of magnitude better than existing server-farm solutions. From an economic standpoint, in addition to introducing reliability problems, systems with large-power envelopes can also be quite expensive from simply an electrical-cost perspective. For example, suppose you operate a data center with 100 Green Destiny racks (a la Google or Yahoo) where each Green Destiny rack consists of 240 processors. At load with disks, each rack consumes 5.2 kW. So,  $5.2 \text{ kW/rack} * 100 \text{ racks} * 24 \text{ hours} * 365 \text{ days} = 4,555,200 \text{ kWh}$  per year. The same number of racks based on a traditional processor could consume  $31.2 \text{ kW/rack} * 100 \text{ racks} * 24 \text{ hours} * 365 \text{ days} = 27,331,200 \text{ kWh}$ . At \$0.15 per kilowatt hour (kWh) in California, this translates to an annual difference of \$3.4 million in energy costs alone.

This isn’t the complete story, however. Because of high heat dissipation, the system with traditional processors must be specially cooled with roughly the same amount of power, for a total power consumption of 54,662,400 kWh, resulting in a total difference of \$7.5 million annually! Not only could California save a lot of money by adopting power-aware (super)computing, but perhaps it could have even avoided the rolling California blackouts during the summers of 2000 and 2001.□

## LOVE IT, HATE IT? LET US KNOW:

feedback@acmqueue.com or [www.acmqueue.com/forums](http://www.acmqueue.com/forums)

## REFERENCES

1. Moore, G. Cramming more components onto integrated circuits, *Electronics* 38, 8 (April 1965).
2. When calculating the price/performance ratio, another metric that is sometimes used in conjunction with the performance metric, price is defined to be the cost of acquisition only and does not account for the cost of operation.
3. Bell, G. and Gray, J. What’s next in high-performance computing? *Communications of the ACM*, 45, 2 (Feb. 2002).
4. Sterling, T., Becker, D., Savarese, D., Dorband, J., Ranawake, U., and Packer, C. Beowulf: A parallel workstation for scientific computation, *Proceedings of*

# making a case for Efficient Supercomputing

- the International Conference on Parallel Processing* (August 1995).
- Feng, W., Warren, M., and Weigle, E. The bladed Beowulf: A cost-effective alternative to traditional Beowulfs, *Proceedings of IEEE Cluster 2002* (Sept. 2002).
  - Warren, M., Weigle, E., and Feng, W. High-density computing: A 240-processor Beowulf in one cubic meter, *Proceedings of Supercomputing 2002* (Nov. 2002).
  - Darling, A., Carey, L., Feng, W. The design, implementation, and evaluation of mpiBLAST, Best Paper: Applications Track, *Proceedings of ClusterWorld Conference & Expo* (June 2003).
  - Karp, A. Speeding up n-body calculations on machines lacking a hardware square root, *Scientific Programming*, 1, 2 (1992).
  - The performance of ASCI Q is extrapolated from the measured performance on a smaller version of the machine with the same architecture. The extrapolation is optimistic; actual performance will likely be somewhat smaller. The power and space numbers for Avalon and Green Destiny are actual measurements, whereas the power and space numbers for the ASCI machines are based on personal communications with system administrators and quoted numbers from the World Wide Web.
  - LANL researchers outfit the “Toyota Camry” of supercomputing for bioinformatics tasks, *BioInform/GenomeWeb* (Feb. 3, 2003).
  - Warren, M., Germann, T., Lomdahl, P., Beazley, D., and Salmon, J. Avalon: An Alpha/Linux cluster achieves 10 Gflops for \$150K, *Proceedings of Supercomputing 1998 (SC'98)* (Nov. 1998).
  - If Green Destiny+ had been specified in a full configuration—that is, 1.125 GB of memory per node and 160 GB of disk per node—the memory density and disk density would have increased by an order of magnitude to 187,500 MB per square foot and 6,400 GB per square foot, respectively. These numbers will have tremendous implications to Web-server farms and search-engine farms like Yahoo and Google.
  - LINPACK, the benchmark used to rank supercomputers in the Top 500 Supercomputer List (<http://www.top500.org>), was chosen because it is the only common benchmark result that we have that has been run across the two different machines—Green Destiny and Japanese Earth Simulator.
  - We note again that if Green Destiny+ had been specified in a full configuration—1.125 GB of memory per node and 160 GB of disk per node—the memory density and disk density would have increased by an order of magnitude to 187,500 MB per square foot and 6,400 GB per square foot, respectively.
  - The Japanese Earth Simulator actually occupies two floors, each 50 meters by 60 meters (or 35,145 square feet) in dimension. Thus, its footprint is effectively  $2 * 35,145 = 70,290$  square feet.
  - Bell, Gordon. Letter to Los Alamos National Laboratory, 2003
  - Lakhman, K. Craig Venter goes shopping for bioinformatics to fill his new sequencing center, GenomeWeb, Oct. 16, 2002; <http://www.genomeweb.com/articles/view-article.asp?Article=2002101693617>.

**WU-CHUN FENG** is a technical staff member and team leader of research and development in advanced network technology (RADIANT) in the Computer and Computational Sciences Division at Los Alamos National Laboratory (LANL). He is also a fellow of the Los Alamos Computer Science Institute and the founder and director of the Advanced Summer Curriculum for Emerging Network Technologies (ASCENT). Feng joined LANL in 1998, where he has been conducting research in high-performance networking and computing. Since then he has established a respected record of over 50 journal and conference publications and has given over 20 invited talks and colloquia. His team at Los Alamos, in conjunction with Caltech, CERN, and SLAC, also recently broke the Internet2 Land Speed Record with a single TCP/IP stream.

Feng received a B.S. in electrical and computer engineering and music (Honors) and an M.S. in computer engineering from the Pennsylvania State University in 1988 and 1990, respectively. He earned a Ph.D. in computer science from the University of Illinois at Urbana-Champaign in 1996.

© 2003 ACM 1542-7730/03/1000 \$5.00.