

POWER AND COOLING IN THE DATA CENTER

Addressing today's and tomorrow's challenges with the AMD Opteron™ processor and AMD PowerNow!™ technology with Optimized Power Management (OPM)

Today's data center isn't merely a room built to house rack upon rack of servers—it's a system. In fact, the data center has become so complex—with so many interactions among processors, rack systems, power systems, data storage, cooling systems, networks and communications channels—that the savvy data center manager has come to regard it as an ecosystem. As such, it requires only the input of energy to become a world unto itself, yet it supports virtually all of the critical business and scientific computing needs both inside and outside the enterprise.

The escalating speed, capacity and power density of the data center's components have increased the interdependence for maximum performance and cost-effectiveness. Processor power, semiconductor density, and the resulting heat generated and power consumed have a profound effect on the cost of running a data center. In these budget-conscious times—in which IT leaders are under unprecedented pressure to manage costs and TCO—it's no longer acceptable to let data center costs spiral out of control.

AMD has brought several of its technologies to bear on the problem, and has capitalized on its traditional strengths in low-power, high-performance processors, integration and power management. Its marriage of AMD Opteron™ processors and AMD PowerNow!™ technology with Optimized Power Management (OPM) provides a fresh set of choices for information management at a critical time in the growth of data centers. This paper examines the underlying technologies and how they combine to achieve a previously unattainable locus of performance, density and low operating cost.

Joining Forces

The AMD Opteron processor has a well-deserved reputation for high performance and low power consumption. Its industry-leading dynamic power management solution, AMD PowerNow! technology, first became available in June 2000. In fact, AMD was the first company to introduce dynamic frequency and voltage support (P-states) capability in x86-based processors. This technology, which has been continually refined, delivers performance on demand and greatly

TABLE OF CONTENTS

Joining Forces	1
Power Challenge and TCO	3
Designed-In Efficiency	4
Rack Density	5
Technical Details of AMD PowerNow!™ Technology with OPM	5
Measurable Savings	6
Additional Power Management Benefits	7
The Dual-Processor Software Bargain	8

reduces power consumption when full CPU performance is not needed. In its latest, most advanced form, AMD PowerNow!™ technology with OPM allows the processor to run at multiple frequencies and voltages without changing the memory/front-side bus speed, under industry-standard ACPI (Advanced Configuration and Power Interface) program control.

Coupled with AMD's pioneering dual-core technology—which puts two full 64-bit CPUs onto one die in every package—AMD PowerNow! technology creates an unparalleled value and performance proposition for the data center manager. As will be discussed later (see “Designed-In Efficiency,” page 4), the AMD Opteron™ processor was designed from its inception to be multi-core capable, and the transition from single-core to dual-core processors can be simple and transparent, with nothing more than a BIOS upgrade. In addition, AMD dual-core technology operates within the same power envelope as single-core, delivering optimal performance per watt.

Optimizing CPU power for the load creates a ripple effect of benefits in the data center that increase dramatically as the number of processors rises. The amount of electricity required to run the processors becomes significant, as does the power to cool them and their power supplies. By offering greater granularity and lower overall power consumption, AMD makes it possible to build data centers with greater density and lower utility cost without expensive retrofitting of higher-capacity cooling systems.

Power Challenge and TCO

Data center managers don't have the option of rebuilding their infrastructure each time rack density or power consumption increases. In fact, data centers are often designed to have a 20-year lifespan.¹ AMD understands this as a responsibility to design and manufacture processors with lower power consumption and waste-heat generation. Many existing data centers were designed around a heat load of 40 to 70 watts per square foot.

However, in today's data center environments, the highest-density racks can exceed 200 watts per square foot, so designers are specifying new data centers to handle heat loads of 350 and even 500 watts per square foot.

Power consumption for cooling has long been regarded as a fixed cost. However, intelligent cooling system design, coupled with AMD PowerNow! technology-enabled systems that can adjust their power consumption based on the performance demands of the workload, can make cooling power a variable, manageable cost. The raw costs of electrical power to run the facility can likewise be managed, because the processors consume a significant percentage of total system power.

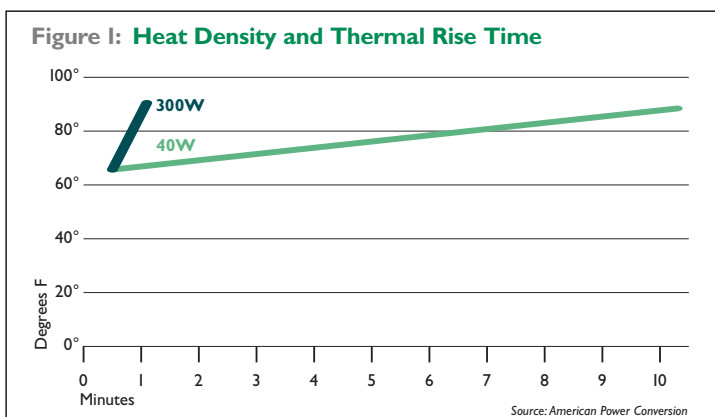
Maximizing the utilization of existing data center facilities is critically important to keeping cost of ownership from spiraling out of control. While typical racks installed in data centers just two years ago might have consumed two kilowatts and emitted 40 watts of heat per square foot, new, high-density racks will consume 10, 15, or even 25KW per rack and may dissipate as much as 500 watts per square foot by the end of the decade. Construction costs, which include cooling, could skyrocket. While it costs roughly \$400 per square foot to build a data center that can handle 40 watts per square foot, 500 watts per square foot may actually cost as much as \$5,000 per square foot for the necessary air conditioning, power supply, power conditioning, and other equipment. Thus a 50,000-square-foot facility might have cost \$20 million a few years ago, but building for the projected future heat load could cost as much as \$250 million today.²

Cooling is the chief driver of increased data center construction cost. With the increased power density of modern rack systems, thermal rise is much faster than in older, lower-density systems. While a data center with an average heat load of 40 watts per square foot can cause a thermal rise of 25 degrees F in 10 minutes, a load of 300 watts per square foot can cause the same rise in less than a minute (see Figure 1).

The startling rapidity of these rise times can mean two things:

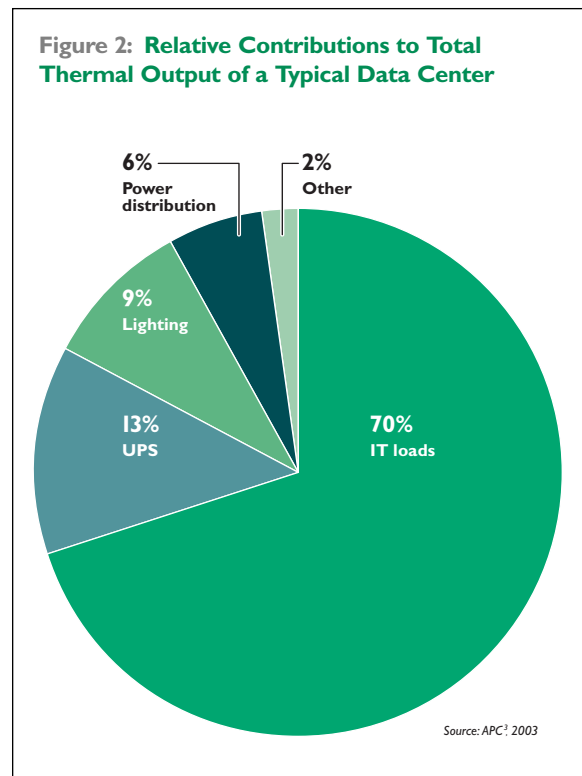
1. A data center with such high-density racks requires an uninterruptible cooling system.
2. Data processing equipment must be highly responsive to thermal rise, and able to regulate itself and reduce operating temperatures in an orderly fashion if cooling should fail or become impaired.

Given the heat density of modern rack systems, redundant, uninterruptible cooling systems are becoming increasingly popular. Moving from the prevalent n+1 style of cooling system back up to a 2n+1 configuration is obviously expensive, but it's absolutely necessary in order to manage



the thermal load generated by today's high-performance servers (see Figure 2 for thermal output by system component).

The data illustrates the fact that anything that can be done to reduce heat and increase computational density is a wise investment. It is significantly less expensive to replace blades and racks with systems featuring cooler-running, energy-efficient processors than to build new facilities to house unnecessarily hot components.



Designed-In Efficiency

Even if they were not equipped with AMD PowerNow!™ technology, the current offerings in the AMD Opteron™ processor product line would be excellent choices for high-density server applications. For example, the standard power, dual-core AMD Opteron processor draws only 95 watts maximum, which includes an integrated memory controller. AMD PowerNow! technology enhances the energy consumption performance of these devices with multiple levels of lower clock speed and voltage, reducing power consumption by as much as 43 percent while under an approximate 60 percent load and as much as 75 percent during idle times.

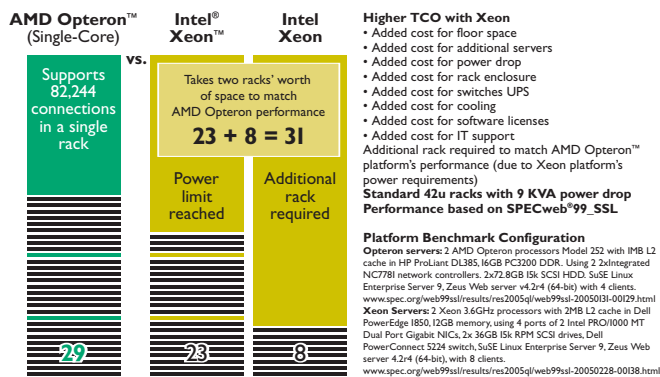
The AMD Opteron processor was indeed designed from the start to be a multi-core design. Having two identical processor cores on the same chip—not just in the same package—provides a number of important benefits:

1. The “edge” components—drivers, voltage regulators, bus interfaces and the like—are among the most power-hungry. If they can be shared among two logic cores, the advantages in terms of lowering power consumption are obvious.
2. Chip-level interconnects are inherently faster than chip-to-chip interfaces, yielding a significant performance advantage. The on-chip system request queue that mediates traffic to and from the CPU cores is an integral part of the overall design.
3. AMD PowerNow! technology P-states are automatically synchronized between CPU cores so that the workload may be shared transparently, giving the highest throughput at the lowest power.
4. Dual-core AMD Opteron processors use the same socket as the single-core version and run in most systems with nothing more than a BIOS upgrade. They were also designed to fall within the same thermal envelope as single-core AMD Opteron processors, but with higher performance, resulting in better performance per watt.

Direct Connect Architecture addresses and helps reduce the real challenges and bottlenecks of the system. Having the CPUs directly connected to CPUs, CPUs directly connected to I/O, and memory directly connected to the CPUs helps reduce chip count, lowering power consumption. The integrated memory controller in AMD Opteron processors effectively eliminates the traditional “north bridge” chip. This is quite beneficial to the system overall, since every piece of silicon on a server circuit board costs money and adds heat.

To the practiced eye, the absence of a separate memory controller chip on the motherboard is obvious. The savings in board real estate, power consumption, and overall system cost is illustratable, and can be quantified. Again, by integrating the memory controller into the same chip that holds the two CPUs in dual-core AMD Opteron processors, AMD has eliminated bus interface chips, synchronizing signals, and similar overhead, saving tens of watts per system. The accompanying heat load is similarly lower. The benefits add up very quickly, particularly in rack systems.

Figure 3: AMD Opteron™ Processor Power Advantage: Increasing Density Without Facilities Upgrades



Rack Density

By offering the AMD Opteron™ processor with AMD PowerNow!™ technology with OPM in both low-power and standard-power versions, AMD offers data center managers unprecedented flexibility to optimally configure rack density, number of processors, and processor performance. For situations where power consumption is more important than having the fastest CPUs, the AMD Opteron HE processor offers all the benefits of its standard-power siblings, including dual-core, 64-bit instruction set, and the same AMD PowerNow! technology with OPM capabilities. But at only 55 watts maximum consumption, it permits excellent rack density while delivering very respectable performance. Like all AMD Opteron processors, the AMD Opteron HE processor can run with up to 75 percent power savings at idle.

In one example, a publicly available power calculator program was used to configure a 48u rack with a power drop of 9KVa. The power calculator takes into account the number of processors, their speed, the amount of RAM, number of PCI cards, and number and capacity of disk drives. Using a well-known, competing workstation/server CPU, the estimated maximum number that could be installed within the rack's thermal/power envelope was 23 lu blades. With all other factors remaining the same, the rack could accept 29 AMD Opteron processor-powered blades, or a 25 percent increase (see Figure 3).

The above numbers are estimates and any such comparisons need to be made in the context of a given data center's total power and thermal envelope. But the benefits are clear and substantial: the amount of computing power per watt, per square foot, and per BTU

can be invariably higher when using the AMD Opteron processor with PowerNow! technology with OPM.

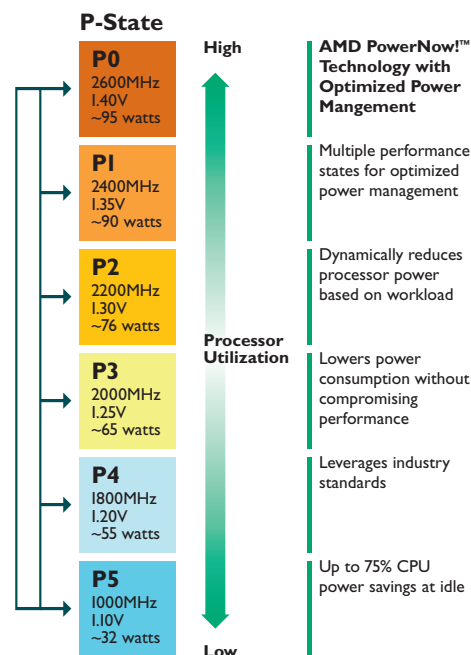
Technical Details of AMD PowerNow!™ Technology with OPM

AMD was first to include dynamic frequency and voltage support (P-states) in x86 processors back in 2000 with the AMD K6®-2 processor. Although AMD PowerNow! is a technological breakthrough and one of the most sophisticated processor power management solutions available, it is implemented with industry-standard ACPI calls.

ACPI was originally developed for power management on notebook computers, but in its latest iteration it has become an operating system-independent power management scheme with inherent multiprocessor support that can tailor each processor's power consumption level to its workload when the platform supports multiple P-states.

AMD PowerNow! technology establishes multiple combinations of P-states (see Figure 4). Although these states can be defined and managed in the BIOS, a multiprocessor system will implement them in the operating system's kernel. The AMD Opteron processor can respond very quickly, in microseconds, to P-state

Figure 4: AMD Opteron™ Processor P-State Support



Source: AMD, 2005

changes so that processor power and speed closely follow the workload. As a result, heat doesn't build up unnecessarily, even in extremely discontinuous workload scenarios.

The highest P-state runs the processor at full clock speed and full voltage. But during off-peak conditions, the clock can drop all the way back to a 1GHz "idle," saving as much as 75 percent of the full-speed power. Granularity is the key to minimizing power consumption: Just as jet fighters in formation use terrain-following radar to "hug" the ground, AMD PowerNow!™ technology with Optimized Power Management (OPM) continually adjusts the performance of each processor, where a mere High, Medium, and Low scheme simply will not provide the optimal balance of maximum performance and power consumption. Although ACPI was originally developed by several companies in conjunction with Microsoft, it is platform-independent. The calls are also implemented in other operating systems, including Linux, and are expected in future Solaris versions. This means that the benefits of AMD PowerNow! technology can be delivered across solutions on any AMD Opteron™ processor-based platform.

Microsoft Windows® XP has a number of power-down states, including global level, sleep states, device-level states, and four dedicated C-states for the CPU. Within its highest-power, or normal, state (C0) the processor driver can impose any number of subdivided "performance states." AMD PowerNow! is implemented in this driver, and it operates independently of Windows® XP as long as Windows® has invoked the C0 state.

While it is somewhat unusual for server operating systems to invoke sleep states, if the OS were to go into standby it would instruct the CPU to go into its C3, or lowest-power, state. By implementing a processor driver, AMD provides that data center managers have the means to fully harness the capabilities of AMD PowerNow! with any operating system, and they have the potential for interoperation with cooling systems, power systems, and facilities management software.

Windows® XP organizes its power management directives into Policy and Non-Policy groups, which communicate to the AMD Opteron processors through ACPI. These directives map well to AMD PowerNow! technology's capabilities. For example, the adaptive policy (an automatic mode in which P-states adjust dynamically depending on workload) will drop the performance of the processor

to the lowest voltage/frequency state available whenever there is not enough demand on the processor to justify the use of a higher state. The adaptive policy will not utilize linear stop-clock throttle states, except in response to thermal/over-temperature events.

Non-Policy states, by contrast, are exceptions that, within the scope of server applications, conserve power in emergencies or prevent thermal damage. If the temperature exceeds a passive thermal trip point, stored as a registry value, the operating system will use successively lower performance levels to reduce temperature to below the trip point. If lower performance levels don't alleviate the problem, the kernel uses stop-clock throttling to prevent damage.

Linux also has proactive Operating System Power Management (OSPM) cooling policies⁴ and a full implementation of ACPI, with table-driven settings that map to AMD PowerNow! P-states. Although OSPM can poll thermal zones, most implementations use asynchronous signaling to eliminate wasteful polling overhead. Like Windows®, OSPM encompasses both active (for performance) and passive (for energy conservation) cooling models and implements critical trip points with multiple thermal thresholds for system protection and orderly idling or shutdown in the case of thermal failure.

Solaris provides a scripted, command-line interface control language that affords data center managers full command over the OS-independent ACPI and lower-level P-states, as well as other operating conditions.⁵ Thermal management can be tailored very closely to the data center's needs.

Measurable Savings

Benchmarks are not your applications, but they are valuable predictors of how your applications will run and how application-like loads affect the power consumption of AMD Opteron processors with AMD PowerNow! technology. AMD has tested the new dual-core AMD Opteron processor in a variety of scenarios, using SPECweb and SPECcpu, measuring power consumption at the processor and at the line cord.

Figure 5 demonstrates how AMD PowerNow! minimizes CPU power consumption as load decreases, and how Optimized Power Management maximizes the benefit. The savings from OPM range from 7 percent at 80 percent CPU utilization to 73 percent at 20 percent CPU utilization.

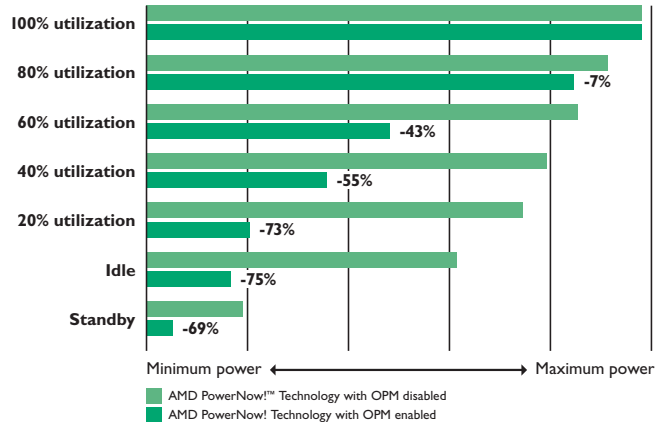


Although the processors are just a fraction of a server's total power budget, the effect is still measurable at the line cord. The savings are most dramatic in the 60 through 20 percent utilization levels, which is where many server processors operate, with intermittently higher loads. When the difference is multiplied out across all servers in a data center, the savings become substantial (see Figure 6).

Figure 5: CPU Power Analysis

AMD PowerNow!™ technology can provide up to 75% power savings at the CPU level with no measurable performance impact.

Average CPU Core Power (SPECWebSSL)



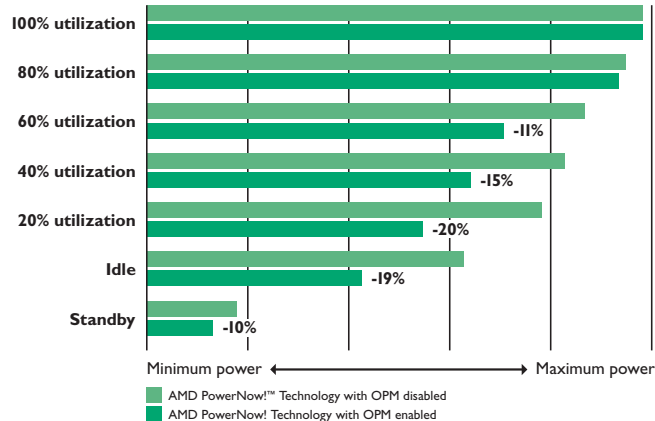
AMD recognizes two other potential opportunities for additional power savings:

1. More efficient power supplies that better translate lower on-board consumption-to-line-cord savings.
2. Cooling systems that more closely follow the system's thermal characteristics.

Figure 6: System Power Analysis

AMD PowerNow!™ technology can provide up to 20% total system power savings with no measurable impact.

Average System Power (SPECWebSSL)

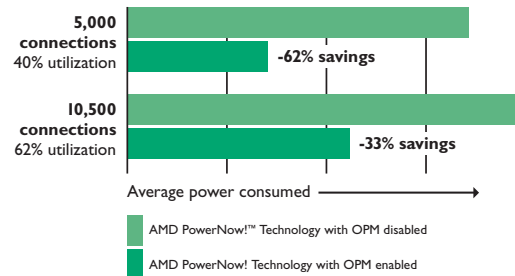


For remote-access/Web server applications, SPECweb is highly relevant because it establishes a benchmark workload typical of Web servers and then measures the number of connections that the system under test can support. The CPU power consumption is measured at a given workload while AMD PowerNow!™ technology is enabled and disabled. The results clearly demonstrate that AMD PowerNow! technology's granularity can give substantial power savings even while maintaining high workloads.

Even in a high-utilization scenario, with 10,500 connections and 62 percent CPU utilization, enabling AMD PowerNow!

Figure 7: Average CPU Core Power (SPECweb)

Average CPU Core Power (SPECWeb)



resulted in a 33 percent drop in CPU power consumed (see Figure 7). With 5,000 connections, which resulted in 40 percent CPU utilization, AMD PowerNow! with OPM was able to reduce CPU power consumption by 62 percent. When the system under test was idle, AMD PowerNow! reduced consumption by an impressive 75 percent. It's important to remember that the cost of electricity is just the tip of the iceberg in such scenarios; all power- and heat-related operating costs drop too.

Additional Power Management Benefits

AMD PowerNow! technology opens a whole new range of increased uptime under the most adverse conditions with the capability to deal with thermal failures without shutting down CPUs. For example, ACPI calls can deal with a failed chiller by throttling back all CPUs to minimal heat dissipation, keeping the system alive with scaled-back performance until the problem can be resolved. Likewise, a local failure in a supplemental cooler or exhaust fan may cause a given rack or blade to run hot. The individual blade or the entire rack can be set to an appropriately low P-state until cooling can be restored.

Cool-aisle/hot-aisle layouts have proven to be highly efficient in maximizing the effectiveness of cooling systems.⁶

THE DUAL-PROCESSOR SOFTWARE BARGAIN

Greater demand on data centers, higher processing requirements, increasing rack density, higher heat loads, energy consumption, and the cost of new and upgraded data center cooling equipment are on a convergent path paved with capital costs. Data center managers are under greater pressure than ever before to make wise investments in new technology. The benefits of these investments are:

1. Higher computational power per watt
2. Lower heat dissipation per unit of computational power
3. Highly granular power management for minimum consumption and excess heat
4. Industry-standard interfaces to power management for easy integration into total data center management
5. Dual-core processors can increase performance by 70 to 90 percent over equivalent single-core processors

Given the other benefits of dual-core AMD Opteron™ processors, it's possible to inadvertently overlook this last point: nearly twice the processing power at no increase, perhaps even a reduction, in heat load. Also, pricing trends in software licensed for dual-core processors are very favorable: more processing power without an attendant increase in per-processor licensing. Microsoft started the trend when it announced in late 2004 that it would maintain the same prices for dual-core chips as for single-core processors.⁷

IBM is one of the first major applications vendors to have recognized the dual-core AMD Opteron processor as one processor, even though it has two fully functional CPUs.⁸ Assuming that others follow this lead, upgrading from single-core to dual-core processors should not affect the cost of software that is licensed on a per-processor basis.

AMD has leveraged its expertise in low-power, highly efficient processors to produce the most advanced, power-efficient family of processors. With its equally innovative AMD PowerNow!™ technology, you can see from the foregoing that the dual-core AMD Opteron processor is the processor of choice for the data center.

Data center managers should seriously consider leveraging this technique if their racks are laid out in a more haphazard manner or if hot-air exhaust is not uniformly oriented towards the rear of the racks. Cool-aisle/hot-aisle implementations are also effective for spot cooling with auxiliary fans or a small refrigeration unit to direct additional cold air to a particular piece of hot-running equipment. Retrofitting with modern racks and blades offers multiple benefits:

1. Preserved investment in existing cooling and power conditioning infrastructure.
2. Decreased cooling power consumption.
3. Increased processor density.
4. Decreased power consumption per blade.

In the case of power failures, an ACPI interface with the UPS can maximize battery life or standby-generator fuel consumption by using AMD PowerNow!™ technology with OPM to reduce the speed and voltage of some or all of the processors. Critical operations can be maintained at full speed, while others can be throttled back. While full integration with cooling and power management systems requires full participation of the rack, power and cooling vendors, all of the hardware and software interfaces are defined and operational today. Facilities management software is currently being adapted to the AMD Opteron™ processor's unique capabilities.

Planning for the unplanned is a key part of the data center manager's job. Every IT manager knows, however, that on a day-in, day-out basis, peak workloads and idle times become predictable. AMD PowerNow! technology with OPM is ideally suited for matching low-load times with low-power P-states, and thanks to the granularity of AMD PowerNow!, maximum power savings are achievable.

1 "Blade Servers and Beyond: Adaptive Cooling for the Next Generation of IT Systems," Liebert Corp., 2005

2 "Data Centers Get a Makeover," Computerworld, November 1, 2004

3 "Calculating Total Cooling Requirements for Data Centers," American Power Conversion, 2003

4 SourceForge.net, 2005

5 Interview, John Fowler, EVP Network Systems Group, Sun Microsystems

6 "Alternating Cold and Hot Aisles Provides More Reliable Cooling for Server Farms," The Uptime Institute, 2002

7 "Microsoft Reveals Licensing Plans for Dual-Core Chips," eWeek, October 19, 2004

8 "IBM Changes Dual-Core Pricing for New AMD Chips," Computerworld, April 22, 2005