

# Managing Data Center Power and Cooling

## with AMD Opteron Processors and AMD PowerNow! Technology

Avoiding unnecessary energy use in enterprise data centers can be critical for success. This article discusses the power and cooling advantages of AMD™ Opteron™ processors and AMD PowerNow!™ technology with Optimized Power Management, which are available in Dell™ PowerEdge™ servers.

BY BRENT KERBY

### Related Categories:

AMD

AMD Opteron processors

Dell PowerEdge servers

Power and cooling

Power management

Visit [www.dell.com/powersolutions](http://www.dell.com/powersolutions) for the complete category index.

Today's data center is more than a room built to house rack upon rack of servers—it is a complex system. In fact, data centers have become so complex—with so many interactions among processors, rack systems, power and cooling systems, storage arrays, networks, and communications channels—that they can be regarded as ecosystems, requiring only the input of energy to become a world in themselves, while supporting virtually all critical business and scientific computing needs both inside and outside an enterprise.

The escalating speed, capacity, and power density of data center components have increased the degree of interdependence typically required for maximum performance and cost-effectiveness. Processor power, chip density, and the resulting power consumption and heat can have a profound effect on data center operational costs. And with data center managers under pressure to

contain budgets and total cost of ownership, managing these effects can be critical to enterprise success.

AMD has developed several technologies to address this issue and has capitalized on its traditional strengths in low-power, high-performance processors; integration; and power management. The result: AMD Opteron processors and AMD PowerNow! technology with Optimized Power Management (OPM), which combine high-efficiency processing with sophisticated granular power management at the processor level. This article examines the underlying technologies of these products and how combining them can enhance efficiency and performance to help reduce data center costs.

### Power and cooling challenges

Power consumption for cooling has long been regarded as a fixed cost. However, intelligent cooling system design,

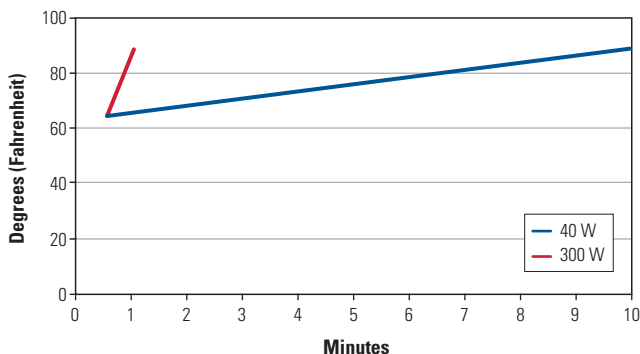


Figure 1. Heat density and thermal rise time

coupled with systems that can adjust power consumption based on performance demands, can make cooling power a variable, manageable cost that avoids wasted energy. The raw costs of electrical power to run a data center facility can likewise be managed, because processors typically consume a significant percentage of total system power.

AMD takes these requirements as a responsibility to design and manufacture processors with the lowest possible power consumption and waste-heat generation. Maximizing existing data center facilities is critical to controlling total cost of ownership—data center managers cannot simply rebuild their infrastructure every time rack density or power consumption increases. Although typical racks installed in data centers just a few years ago might have consumed 2 kW of power and emitted 40 W of heat per square foot, in the coming years high-density racks may consume up to 25 kW per rack and dissipate as much as 500 W per square foot. This increased power density can cause much faster thermal rise than in older, lower-density systems: a data center with an average heat load of 40 W per square foot can cause a thermal rise of 25°F in 10 minutes, while an average heat load of 300 W per square foot can cause the same rise in less than a minute (see Figure 1).<sup>1</sup>

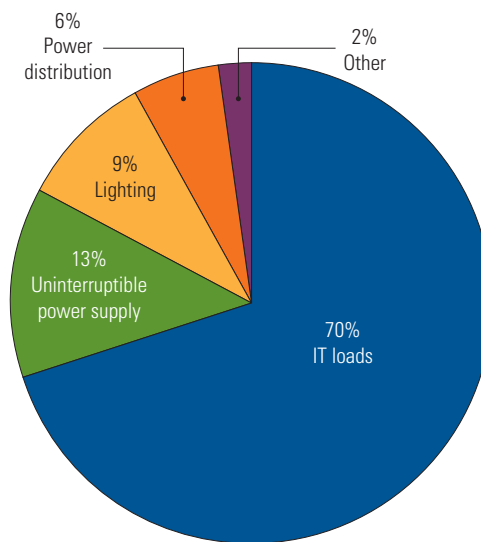
And because cooling is a major factor in data center construction costs, this type of increase can cause overall construction costs to skyrocket. For example, in 2004 it might have cost roughly US\$400 per square foot to build a data center that emitted 40 W per square foot. Today, high-density racks producing up to 500 W per square foot can cost as much as US\$5,000 per square foot for the necessary air conditioning, power supply, power conditioning, and other equipment. Thus a typical 50,000-square-foot facility that might have cost US\$20 million a few years ago could cost as much as US\$250 million today when factoring in projected future heat loads.<sup>2</sup>

Figure 2 shows where heat comes from in a typical data center. Because the majority of heat is generated by IT loads, reducing

heat from these loads is a significant part of managing data center cooling in a cost-effective way: it can be significantly less expensive to replace blades and racks with reduced-heat, energy-efficient processors than to build new facilities to house unnecessarily hot components. To handle increased thermal rise times, data centers with high-density racks typically should include an uninterruptible cooling system; in addition, data-processing equipment must be self-regulating, highly responsive to thermal rise, and capable of reducing operating temperatures in a controlled way if the cooling system fails or becomes impaired. Given the heat density of modern rack systems, redundant, uninterruptible cooling systems are becoming increasingly popular. Changing a redundant cooling system from the prevalent  $n + 1$  configuration to a  $2n + 1$  configuration incurs costs, but it may be necessary to help reduce heat in the data center effectively.

**AMD Opteron processors with AMD PowerNow! technology**

AMD Opteron processors are designed to offer high performance and low power consumption. Their AMD PowerNow! technology with OPM uses dynamic frequency and voltage support, and has been refined since its introduction in June 2000 to deliver performance on demand while greatly reducing power consumption when full processor performance is unnecessary. In its latest, most advanced form, AMD PowerNow! technology with OPM under industry-standard Advanced Configuration and Power Interface (ACPI) program control allows processors to run at



Source: "Calculating Total Cooling Requirements for Data Centers," by American Power Conversion, www.apcmedia.com/salestools/NRAN-5TE6HE\_R1\_EN.pdf.

Figure 2. Relative contributions to total thermal output of a typical data center

<sup>1</sup> Bob Sullivan, senior consultant at the Uptime Institute, quoted in "Data Centers Get a Makeover," by Gary Anthes, in *Computerworld*, November 1, 2004, www.computerworld.com/printthis/2004/0,4814,97021,00.html.

<sup>2</sup> Ron Hughes, president of the California Data Center Design Group, quoted in "Data Centers Get a Makeover," by Gary Anthes, in *Computerworld*, November 1, 2004, www.computerworld.com/printthis/2004/0,4814,97021,00.html.

multiple frequencies and voltages without changing memory or frontside bus speeds.

Even without AMD PowerNow! technology, AMD Opteron processors can provide the type of efficient, low-heat processors for high-density server applications that help reduce overall cooling costs—the standard-power dual-core AMD Opteron processor, which includes an integrated memory controller, consumes only 95 W maximum. Adding AMD PowerNow! technology to this processor enhances energy-consumption performance by allowing multiple levels of reduced clock speed and voltage.

Optimizing processing power can provide numerous benefits for enterprise data centers; these benefits can increase exponentially as the number of processors rises, because of the increasing amount of electricity required to run the processors as well as the power to cool them and additional platform components. By enabling increased granularity and reduced overall power consumption, AMD technology allows enterprises to increase data center density without expensive retrofitting of high-capacity cooling systems.

### Efficient design

Dual-core AMD Opteron processors can enhance performance, efficiency, and cost-effectiveness in enterprise data centers. Having multiple identical processor cores on the same chip offers several important advantages:

- **Low power consumption:** The “edge” components of a processor—drivers, voltage regulators, bus interfaces, and the like—use a substantial amount of its power. Sharing these components between two cores uses less power than if they were duplicated in two separate processors.
- **High interconnect speeds:** 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 cores is an integral part of the overall design.
- **High throughput with reduced power consumption:** AMD PowerNow! technology P-states are automatically synchronized between cores, allowing the workload to be shared transparently and providing increased throughput with reduced power consumption compared to separate processors.
- **High performance per watt:** Dual-core AMD Opteron processors use a common socket that is planned to support future quad-core AMD Opteron processors and be able to run in many systems, requiring nothing more than a BIOS upgrade. Quad-core AMD Opteron processors are intended to fall within the same thermal envelope as dual-core AMD Opteron processors while providing increased performance, resulting in increased performance per watt over dual-core AMD Opteron processors.

The Direct Connect Architecture used in AMD Opteron processors connects the processing cores directly to the integrated memory controller, I/O, and other processors, which helps reduce chip count, system bottlenecks, and power consumption. The integrated memory controller effectively eliminates the traditional Northbridge controller and thus provides several benefits, the most important of which is enhanced performance because the processor no longer has to communicate over a frontside bus to its memory controller or I/O. Because the memory controller is an integral part of the chip, there is no bus contention for the controller—the crossbar switching built into the AMD Opteron processor gives both cores easy access to the memory controller.

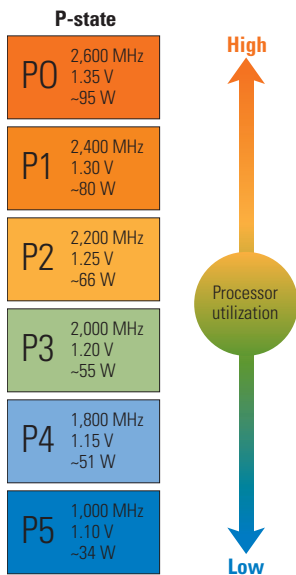
AMD PowerNow! technology provides sophisticated processor power management by establishing multiple combinations of processor voltage and frequency called P-states.

Every piece of silicon on a server circuit board costs money and adds heat. Each additional device that requires a board- or bus-level interconnect is a bottleneck—not a potential bottleneck, but one with a measurable effect on overall system throughput. The savings in board space, power consumption, and overall system cost are clear and quantifiable. By integrating the memory controller into the same chip that holds the two cores in the dual-core AMD Opteron processor, AMD has eliminated bus interface chips, synchronizing signals, and similar overhead, saving tens of watts per system along with the related reduced heat load. These advantages can add up quickly, particularly in rack systems.

And because AMD Opteron processors were designed from their inception to be multi-core capable, a future transition from dual-core to quad-core processors is planned to be simple and transparent, requiring nothing more than a BIOS upgrade. AMD plans for future quad-core AMD Opteron processors to operate within the same power and thermal envelopes as the dual-core models while providing increased performance, helping deliver optimal performance per watt.

### Sophisticated, OS-independent power management

AMD PowerNow! technology provides sophisticated processor power management by establishing multiple combinations of processor voltage and frequency called P-states. Although these states can be defined and managed in the BIOS, multiprocessor systems implement them in the OS kernel. AMD Opteron processors can respond very quickly—in microseconds—to P-state changes, enabling processor power and speed to follow workload closely. The highest P-state runs the processor at full clock speed and full voltage, but during off-peak



Example:  
AMD Opteron 2218 series processor

Figure 3. Power and cooling management with processor performance states

conditions, the clock can drop all the way back to a 1 GHz “idle,” saving as much as 75 percent of the full-speed power, and make continuous granular changes between these two states (see Figure 3). This technology helps avoid unnecessary heat buildup, even in extremely discontinuous workload scenarios.

AMD PowerNow! technology is implemented using the ACPI standard, which was originally developed for notebook power management but has become a general power management scheme with inherent multi-processor support and platform independence. Using industry-standard ACPI calls allows AMD PowerNow! technology to interoperate with cooling systems,

power systems, and facilities management software and with many different operating systems, including Microsoft® Windows® and Linux® operating systems.

**Microsoft Windows XP.** Windows XP has several power-down states, including global-level states, sleep states, device-level states, and four dedicated C-states for the processor. Within its highest-power (or *normal*) C-state, C0, the processor driver can impose several subdivided *performance states*. AMD PowerNow! technology is implemented in this driver, and can operate independently of Windows XP as long as the OS has invoked the C0 state. Although it is somewhat unusual for server operating systems to invoke sleep states, if the OS were to go into standby, it would instruct the processor to go into its C3 state, which uses the lowest amount of power of the C-states.

Windows XP organizes its power management directives into *policy* and *non-policy* groups, which communicate with the AMD Opteron processors through ACPI. These directives map well to AMD PowerNow! technology capabilities. The *Adaptive* policy, for example, reduces processor performance to the lowest voltage and frequency state available whenever processor demand does not justify a higher state. This policy does not utilize linear stop-clock throttle states, except in response to thermal events.

Non-policy states, in contrast, are exceptions that, within the scope of server applications, are designed to conserve power in emergencies or prevent thermal damage. If the temperature exceeds a passive thermal point stored as a registry value, the OS uses successively lower performance levels to reduce the temperature below this point.

If reduced performance levels do not alleviate the problem, the kernel uses stop-clock throttling to help prevent damage.

**Linux.** Linux has proactive OS-directed power management (OSPM) cooling policies and a comprehensive ACPI implementation, with table-driven settings that map to AMD PowerNow! technology P-states. Although OSPM can poll thermal zones, typical implementations use asynchronous signaling to eliminate wasteful polling overhead. Like Windows power management, OSPM encompasses both active (for performance) and passive (for energy conservation) cooling models and implements critical points with multiple thermal thresholds for system protection and orderly idling or shutdown in case of thermal failure.

### Potential benefits of AMD Opteron processors and AMD PowerNow! technology

AMD Opteron processors and AMD PowerNow! technology with OPM can provide the following benefits:

- Optimized computational power per watt
- Reduced heat dissipation per unit of computational power
- Enhanced granular power management to help minimize power consumption and excess heat
- Industry-standard power management interfaces for easy integration into total data center management
- Significant performance increases over equivalent single-core processors

Data center managers should not overlook this last benefit: a dual-core processor typically can provide nearly twice the processing power of a single-core processor without increasing the heat load (and possibly even reducing it). And because some software vendors—Microsoft, for example—maintain the same license prices for dual-core processors as for single-core processors, enterprises can gain increased processing power without attendant increases in licensing costs.

AMD Opteron processors and AMD PowerNow! technology with OPM are designed in particular to provide two benefits: efficient power use and enhanced performance.

### Efficient power use

Data center servers typically have fluctuating workloads throughout the day, causing variable processor utilizations depending on demand. Understanding the average utilization of their installed server base and the power consumption across multiple utilization levels can give data center managers an indication of possible power savings in a particular environment.

Figure 4 illustrates how AMD PowerNow! technology can minimize processor power consumption as load decreases, and how OPM can maximize the potential benefits: AMD PowerNow! technology

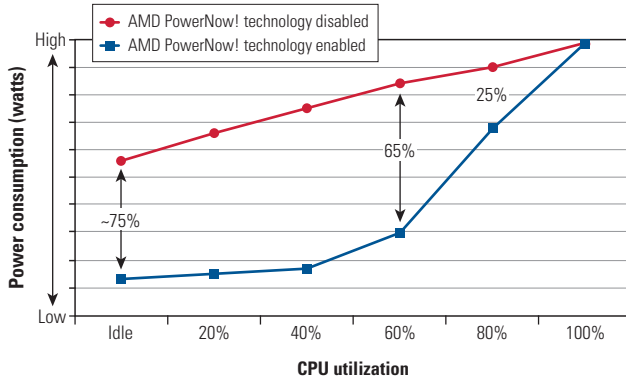


Figure 4. Average CPU core power measured with AMD Synthetic CPU Workload Utility

can provide platform-level power savings ranging from 25 percent at 80 percent processor utilization to 65 percent at 60 percent processor utilization to approximately 75 percent during processor idle.<sup>3</sup>

AMD PowerNow! technology with OPM matches low-load times with low-level P-states to help avoid unnecessary power use. Virtualization, for example, is vitally important in many data centers—and in these environments, virtualized processors, storage volumes, drive spindles, processes, and communication channels are constantly changing. Applications may demand additional resources as workloads increase, but ACPI, which functions at the processor level, can manage power consumption even as conditions change dynamically. Deploying AMD PowerNow! technology with OPM in conjunction with hot aisle/cold aisle layouts<sup>4</sup> and retrofitting with new racks and blades can provide multiple benefits, including allowing data centers to preserve investments in existing power and cooling infrastructures, reducing overall cooling power consumption and power consumption per server, and increasing processor density.


### Enhanced performance

One of the truisms in IT is that there is no such thing as too much bandwidth—and in the server world, today's acceptable bandwidth is tomorrow's bottleneck. AMD has made specific enhancements to the AMD Opteron cache coherency algorithms that are designed to provide each processor extremely rapid and efficient access to other processors' cache contents, even though each processor maintains its own 1 MB level 2 (L2) cache. Because the majority of data center applications are inherently multi-threaded, they can benefit greatly from this architecture. Memory bandwidth can be more important than core clock speed when properly applied, with processors coupled closely through crossbar switching and a memory controller that runs at the processor's speed, as in AMD Opteron processors.

AMD Opteron processors can also provide performance advantages for communicating with memory and other I/O. Because the memory controller is integrated with the chip and runs at the processor's speed, communications between the memory controller and processors avoid frontside bus bottlenecks, which can result in significantly lower latencies than a traditional external memory controller. Multiple AMD Opteron processors communicate with one another through AMD HyperTransport™ links running at 1 GHz, with a total theoretical bandwidth for AMD Opteron 2000 and 8000 series processors of 32 GB/sec. These speeds are critical for processor-to-processor communication, particularly for the AMD Opteron 8000 series processor, which is optimized for four-way and eight-way configurations.

With so many applications executing 32-bit code, it can be tempting to dismiss the 64-bit capabilities of AMD Opteron processors as a “someday” feature. But 64-bit operating systems, compilers, and applications are increasingly available, and depending on the workload, can be dramatically faster than their 32-bit counterparts for compute-intensive tasks. Data center managers should keep in mind that the difference between 32 and 64 bits is not only the number of bits—64-bit systems provide exponentially more memory that can be directly addressed than 32-bit systems.

### Efficient power and cooling in enterprise data centers

For data center managers, the increasing demands, processing requirements, rack density, heat loads, energy consumption, and costs for new and upgraded cooling equipment create greater pressure than ever before to make wise technology investments. Dual-core AMD Opteron processors with AMD PowerNow! technology and Optimized Power Management provide a low-power, high-efficiency option that helps reduce power consumption and cooling costs in enterprise data centers. 

**Brent Kerby** is the product marketing manager for AMD Opteron processors. He has more than 20 years of experience in both technical and marketing positions for consumer and commercial computing products. Brent is currently focused on elements of data center total cost of ownership, including performance-per-watt efficiency and power management.

#### FOR MORE INFORMATION

##### AMD Opteron processors:

[www.amd.com/opteron](http://www.amd.com/opteron)

##### Dell systems featuring AMD processors:

[www.dell.com/amd](http://www.dell.com/amd)

<sup>3</sup> These average results are based on a four-socket internal AMD test platform with four AMD Opteron 8220 SE processors; four 1 GB, 667 MHz double data rate 2 (DDR2) dual in-line memory modules (DIMMs) per socket (for a total of 16 GB of memory); and a 250 GB Serial ATA (SATA) hard drive.

<sup>4</sup> Hot aisle/cold aisle layouts help maximize the efficiency and effectiveness of cooling systems, and lend themselves well to spot-cooling with auxiliary fans or small refrigeration units to direct additional cold air to a particular piece of equipment. For more information, see “Alternating Cold and Hot Aisles Provides More Reliable Cooling for Server Farms,” by the Uptime Institute, [www.uptime.com/TUI/pages/whitepapers/tuiaisles.html](http://www.uptime.com/TUI/pages/whitepapers/tuiaisles.html).