

The Problem of Power Consumption in Servers

ABSTRACT: Capabilities of servers and their power consumption have increased over time. Multiply the power servers consume by the number of servers in use today and power consumption emerges as a significant expense for many companies. The main power consumers in a server are the processors and memory. Server processors are capping and controlling their power usage, but the amount of memory used in a server is growing and with that growth, more power is consumed by memory. In addition, today's power supplies are very inefficient and waste power at the wall socket and when converting AC power to DC. Also, when servers are in operation, the entire chassis will heat up; cooling is required to keep the components at a safe operating temperature, but cooling takes additional power. This article explains how servers consume power, how to estimate power usage, the mechanics of cooling, and other related topics.

Server Power Usage

As data centers and volumes of servers have grown, so has the overall amount of electricity consumed. Electricity used by servers doubled between 2000 and 2005, from 12 billion to 23 billion kilowatt hours. This was due to the increase in the number of servers installed in data centers and to the required cooling equipment and infrastructure (Koomey 2008).

Individual servers are consuming increasing amounts of electricity over time. Before the year 2000, servers on average drew about 50 watts of electricity. By 2008, they were averaging up to 250 watts. As more data centers switch to higher density server form factors, the power consumption will increase at a faster rate. Analysts have forecasted that if the current trend is not abated, then the power to run servers will be equal to or greater than the server cost.

Due to these trends, it is important to understand how a server uses and consumes power. When replacing or upgrading server, it is possible to specify energy efficient improvements.

Power and Server Form Factor

Power use varies with the server's form factor. In the x86 server market, there are four basic server form factors:

- pedestal servers,
- 2U server rack servers,
- 1U rack servers, and
- blade servers.

Where floor space is restricted, and increasing computing capacity is a goal, many data centers utilize rack servers or blade servers rather than pedestal servers.

Servers, routers, and many other data center infrastructure devices are designed to mount in steel racks that are 19 inches wide. For rack servers, the height of the server is stated in multiples of U when 1U equals 1.75 inches. The U value identifies the form factor. Most common are 1U and 2U servers. Power use varies by server form factor due to the individual configuration, the heat and thermal environment related to that configuration, and the workload being processed.

Power and Heat

Much of the electrical energy that goes into a computer gets turned into heat. The amount of heat generated by an integrated circuit is a function of the efficiency of the component's design, the technology used in its manufacturing process, and the frequency and voltage at which the circuits operate. Energy is required to remove heat from a server or from a data center packed with servers.

Computer sub-systems such as memory sub-systems and power supplies, and especially large server components generate vast amounts of heat during operation. This heat must be dissipated to keep the components within their safe operating temperatures. Overheated parts will generally have a shorter maximum life-span. Shorter component lifespan can produce sporadic problems, system freezes, or even system crashes.

In addition to server component heat generation, extra cooling is required when parts of the server are run at higher voltages or frequencies than specified. This is called over-clocking. Over-clocking a server results in increased performance, but also generates a greater amount of heat.

How Cooling is Achieved

Server manufacturers use several methods to cool components. Two common methods are the use of heat sinks to increase the surface area that dissipates the heat and the use of fans to speed up the exchange of air heated by the components for cooler ambient air. In some cases, soft cooling is the method of choice. Computer components can be throttled down to decrease heat generation.

Heat sinks consist of a metal structure with one or more flat surfaces (i.e., a base) to ensure good thermal contact with the components to be cooled, and an array of comb or fin like protrusions. Fins increase the surface area in contact with the air and thus increase the rate of heat dissipation, as Figure 1 illustrates. Heat sinks are frequently used in conjunction with a fan to accelerate airflow over

the heat sink. Fans provide a larger temperature gradient by replacing the warmed air with cool air faster than convection alone can accomplish. Fans are used to make forced air systems, where the amount of air moved to cool components is far greater than the the flow due to convection.

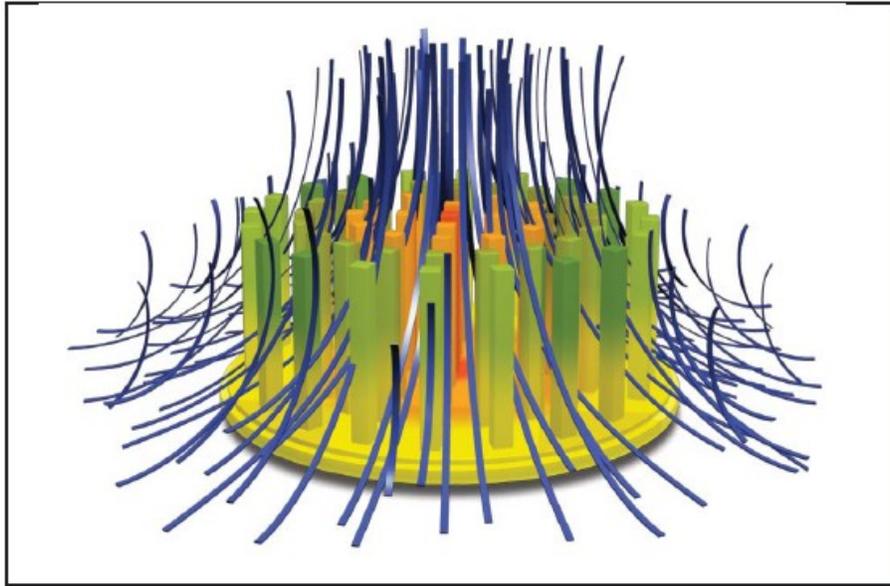


Figure 1 Natural Convection Heat Sink (Source: Wikipedia, 2008)

Heat sink performance is defined as the thermal resistance from junction to case of the component. The units are $^{\circ}\text{C}/\text{W}$. A heat sink rated at $10\text{ }^{\circ}\text{C}/\text{W}$ will get $10\text{ }^{\circ}\text{C}$ hotter than the surrounding air when it dissipates 1 watt of heat. Thus, a heat sink with a low $^{\circ}\text{C}/\text{W}$ value is more efficient than a heat sink with a high $^{\circ}\text{C}/\text{W}$ value.

A quality heat sink can dissipate thermal energy to an extent that additional cooling components need only be minimal. Heat sink thermal performance is determined by:

- *Convection or fin area.* More fins can provide more convection area. But care must be taken if a fan is used with a finned heat sink. In some cases, the pressure drop increases for forced air system.
 - A shortcoming of the conventional fan-top CPU coolers is the reduction of airflow due to pressure drop resulting from the airflow obstruction of the chassis cover and the fins of the heat sink itself.
 - Fan performance is rated in cubic feet per minute (CFM) with 0-pressure drop and performance is severely compromised with only minimal air flow obstructions from either the intake or exhaust side of the fan.

- *Conduction area per fin.* The thicker the fin, the better the conduction heat versus thinner fins.
 - The most energy efficient heat sink designs will balance between multitudes thin fins and fewer thick fins.
- *Heatsink base spreading.* Heat must be spread out as evenly as possible in the base for fins to work effectively. A thicker base is good for heat spreading.
 - However since server form factors are limited to a specific height to fit in racks, the thicker base leads to reduced fin height and hence reduced fin area and increased pressure drop.

Power consumption varies based type of server form factor, and it can also vary with the workload that is processed. Workloads are increasing across all server types due to increases in server processing performance, and thus become another trend increasing power consumption on servers.

Table 1 shows a sampling of power increase by server form factor over time. The three classes of server are defined by IDC.

- Volume servers cost less than \$25,000 and most commonly have one or two processor sockets in 1-2U rack mount form factor.
- Mid-range servers cost between \$25,000 and \$499,999 and are typically contain two to four processor socket or more.
- High-end server cost \$500,000 or more and are typically contain eight processor sockets or more.

Table 1 Estimated Average Power Use (W) per Server, by Server Class, 2000 to 2006 (Source: Koomey J 2007b Estimating Total Power Consumption by Servers in the US and the World. Oakland, CA: Analytics Press)

| Server class | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 |
|--------------|-------|-------|-------|-------|-------|-------|-------|
| Volume | 186 | 193 | 200 | 207 | 213 | 219 | 225 |
| Mid-range | 424 | 457 | 491 | 524 | 574 | 625 | 675 |
| High-end | 5,534 | 5,832 | 6,130 | 6,428 | 6,973 | 7,651 | 8,163 |

A pedestal server varies in width and is designed to optimize the performance and cooling of the server. Because these systems are not space constrained, they have large heat sinks, multiple fans, and great air cooling.

Rack and blade servers are designed to fit within a standardized 19" mounting rack. The rack server architecture and the limited height for air vents and fans make them run hotter, and thus require more power in the data center for cooling infrastructure. 2U servers run hotter than a pedestal server, but cooler than 1 U

servers or blades.

2U server, at 3.5" high can use more and larger fans in addition to bigger heat sinks resulting in improved cooling capability and thus less power consumption than a 1U server. Most servers are designed to bring cool, fresh air in from the bottom front of the case and exhaust warm air from the top rear.

Rack server architecture typically locates customer-desirable features up front, such as disk drives forcing the hot components, such as the server processor and memory, in the back. With rack servers, manufacturers try to achieve a balanced or neutral airflow. This is the most efficient, however many servers end up with a slightly positive airflow which provides the additional benefit of less dust build up if dust filters are used.

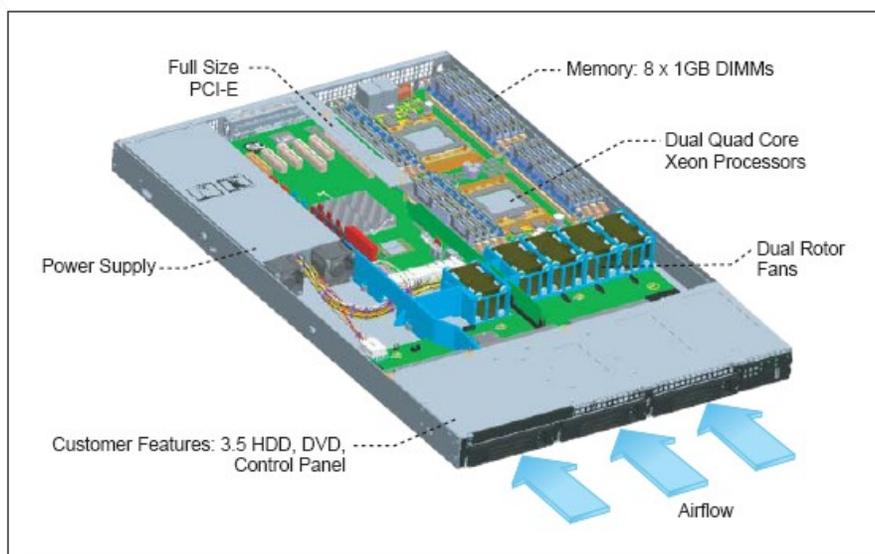


Figure 2 1U Server Architecture (Source: Intel Labs, 2006)

The 1U form factor, shown in Figure 2, and blade server are the most difficult to cool because of density of components and lack of space for airflow cooling. Blade servers have the benefit of more processing power in less rack space and simplified cabling. As many as 60 blade servers can be placed in a standard height 42U rack. However this condensed computing comes with a power price. The typical power demand (power and cooling) for this configuration is more than 4,000 watts compared to a full rack of 1U servers at 2,500 watts. Data centers address this demand by either increasing power consumption or with more exotic ways of computer cooling, such as liquid cooling, Peltier effect heat pumps, heat pipes, or phase change cooling. These more sophisticated cooling techniques all use more power.

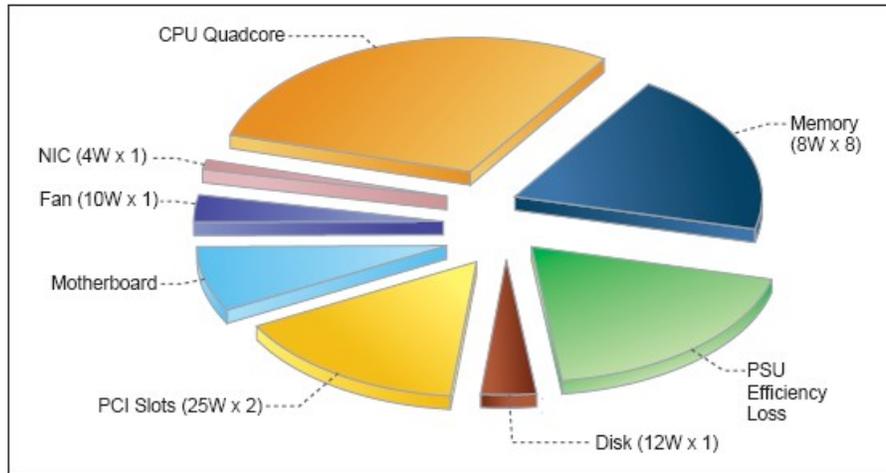


Figure 3 Server Power Consumption (Source: Intel Labs, 2008)

Server Power Breakdown by Components

Figure 3 highlights how power is consumed on average within an individual server. Processors and memory consume the most amount of power, followed by the power supply efficiency loss. Disk drive power only becomes significant in servers with several disk drives.

Processor power consumption varies greatly by the type of server processor used. Power consumption can vary from 45W to 200W per multi-core CPU. Newer Intel processors include power saving technologies, such as Demand Based Switching and Enhanced Speed Step technology. These newer processors also support power saving states such as C1E and CC3. Multi-core processors are much more power efficient than previous generations. Servers using the recent Quadcore Intel® Xeon™ processors can deliver 1.8 teraflops at peak performance using less than 10,000 watts of power. Pentium® processors in 1998 would have consumed about 800,000 watts to achieve the same performance. Server processor power consumption will also vary depending on the server workload.

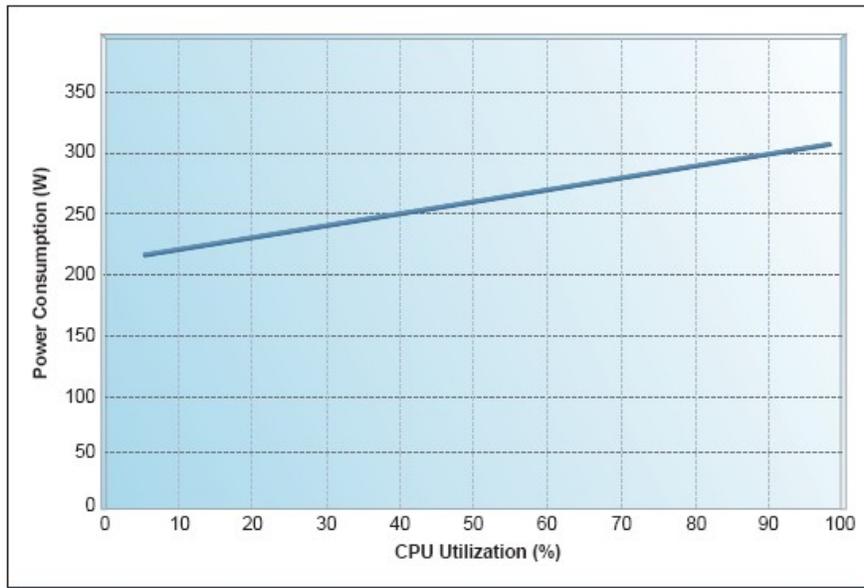


Figure 4 CPU Utilization and Power Consumption (Source: Blackburn 2008)

Figure 4 illustrates how processor energy efficiency (e.g., performance per watt) increases as server utilization increases for a typical workload. Tuning workloads with optimized processor utilization can greatly affect power consumption and energy efficiency.

By taking average processor utilization over a defined period of time, it is possible to calculate an estimate of the power consumed for that period. Many server workloads scale linearly from idle to maximum power. When you know the power consumption of a server at peak usage and at idle it becomes a simple arithmetic operation to estimate power usage at any utilization rate.

Estimating Power Consumption

An estimate of power consumption (P) at any specific processor utilization ($n\%$) can be calculated if power consumption at maximum performance (P_{max}) and at idle (P_{idle}) are known. Use the following formula:

$$P_n = (P_{max} - P_{idle}) \times \frac{n}{100} + P_{idle}$$

For example, if a server has a maximum power draw of 400 watts (W) and an idle power draw of 200W, then at 25 percent utilization the power draw would

approximate to:

$$\begin{aligned} P_{25} &= (400 - 200) \times \frac{25}{100} + 200 \\ &= 200 \times 0.25 + 200 \\ &= 250 \text{ W} \end{aligned}$$

In this example, if the server was running at that average utilization for a 24 hour period, then the energy usage would equate to the following:

$$\begin{aligned} 250 \text{ W} \times 24 \text{ h} &= 600 \text{ Wh} \\ &= 6 \text{ kWh} \end{aligned}$$

Through empirical measurement of various servers using a power meter this approximation has proven to be accurate to within ± 5 percent across all processor utilization rates.

The next largest consumer of power in a server is memory. Intel processor power levels are being well controlled and capped with the latest generations. However, power consumption by memory chips is growing and shows no signs of slowing down in the future. Furthermore, applications continually seek more memory. Here are some of the reasons why demand for memory is growing in servers:

- Increases in processor core counts in latest servers; the more cores, the more memory can be utilized in a server
- Increasing use of virtualization; data centers are adopting virtualization in increasing rates
- In new usages by Internet Protocol Data Centers, such as Google and Facebook, with memory intensive search applications

Memory is packaged in dual in-line memory modules (DIMMs) and these modules can vary in power from 5W up to 21W per DIMM, for DDR3 and FB-DIMM (Fully Buffered DIMM) memory technologies. The memory in a server with eight 1GB DIMMs can easily consume 80W. Many large servers now use 32 and 64 DIMMs, resulting in more power consumption by memory than processors.

For each generation of memory technology, there are key physical and electrical attributes of the DIMM that contribute to its power consumption and bandwidth. The Dynamic Random Access Memory (DRAM) packaging type and die count, number DRAM ranks on a DIMM, data transfer speed and data width define the DIMM capacity and power requirements. DIMMs can have registers, known as RDIMMs or be without registers, known as UDIMMs (Unregistered DIMMs). RDIMMs consume slightly more power than UDIMMs.

Figure 5 shows how power consumption differs among RDIMMs using DDR2 and

DDR3 technology. For power consumption of the latest DIMMs, check the websites of the leading memory manufacturers.

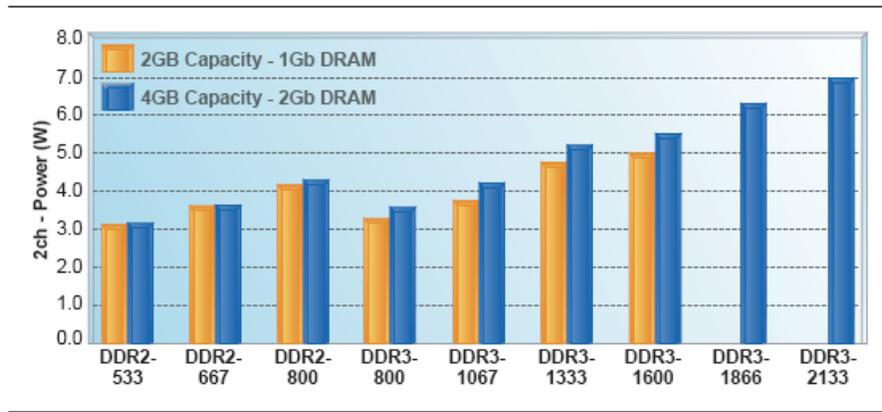


Figure 5 RDIMM Memory Power Comparison (Source: Intel Platform Memory Operation, 2007)

Power consumed by DIMMs is typically measured in Active and Idle Standby states. Active Power is defined as: L0 state, 50 percent DRAM bandwidth, 67 percent read, 33 percent write, with primary and secondary channels are enabled. The DRAM clock is active, and CKE is high. Idle Power is defined as: L0 state; Idle (0 percent bandwidth), Primary channel enabled, Secondary channel disabled, CKE high; Command and address lines stable; and the SDRAM clock active.

On average, DDR3 DIMMs use 5W-12W when active. DIMMs from different vendors will vary based on their manufacturing process for the DRAM components and the components/configuration they use to make the memory module. Additionally, memory power consumed will vary depending upon the application and workload running.

Table 2 shows sample RDIMM power consumptions in 2008, for DDR2 technology running at speeds of 667MHz. Table 2 highlights that power consumption by memory products varies widely among suppliers and configurations. For power consumption of the latest UDIMMs or RDIMMs, check the websites of these vendors.

Table 2 RDIMM Power Consumption by Vendor and Configuration (Sources: Publicly-available datasheets from each vendor, 2008)

| Vendor | Active Power | Idle Power | Configuration |
|---------------|--------------|------------|----------------------|
| 2 GB Capacity | | | |
| Micron | 5W | 1.9W | 2 ranks, 128Mb x 8 |
| Elpida | 6W | 2.2W | 1 rank, 18 1Gb parts |
| Qimonda | 6.2W | 4.0W | 2 ranks |
| Samsung | 6.8W | 2.4W | 2 ranks, 128Mb x 8 |
| Hynix | 7W | 2.4W | 2 ranks |
| Samsung | 10.6W | 2.4W | 1 rank, 256Mb x 4 |
| 4 GB Capacity | | | |
| Hynix | 8.6W | 3.8W | 2 ranks, 256Mb x 4 |
| Qimonda | 8.9W | 5.3W | 2 ranks |
| Micron | 9.3W | 2.6W | 2 ranks, 256Mb x 4 |
| Samsung | 13.1W | 4.2W | 2 ranks, 256Mb x 4 |

As DIMMs increase in capacity, going from 4GB in 2008 to 16GB or 32GB in the near future, their power consumption will increase. DIMMs will also increase in speed over time, which increases the power consumption of the DIMM. Table 3 shows DDR3 RDIMM raw cards, DRAM density, capacity and the forecasted power use based on different speed targets of 1066MHz, 1333MHz and 1600MHz. Power is forecasted to trend higher with the memory speeds of 1866MHz and 2133 MHz. As Table 3 shows, memory power can vary significantly depending upon the memory technology used, the memory configuration, and the vendor.

Table 3 Future DIMM Power Consumption by Frequency, Configuration, and Capacity (Source: Intel Platform Memory Operation, 2008)

| Sample Card | Freq (MHz) | DIMM Configuration | DIMM Tech/Capacity | Power/DIMM | 64GB System Power |
|-------------|------------|--------------------|--------------------|------------|-------------------|
| Card A | 1066 | QRx4 | 2Gb/8GB | 15.5W | 124W |
| Card B | 1333 | QRx8 | 2Gb/8GB | 10.6W | 84.8W |
| Card C | 1333 | DRx4 | 1Gb/4GB | 10.6W | 169.6W |
| Card D | 1333 | QRx4 | 2Gb/16GB | 20.5W | 82W |
| Card E | 1600 | QRx8 | 2Gb/8GB | 10.1W | 80.8W |
| Card F | 1600 | QRx4 | 2Gb/8GB | 19.1W | 152.8W |

Reducing Energy Consumption by Memory Subsystems

Cooling of memory is increasingly challenging and requires additional power in most server systems. In the past, memory bandwidth requirements were sufficiently low that memory was relatively simple to cool and required no thermal enhancements on the DIMM, no thermal sensors, nor throttling. The opposite is now true. The thermal analysis of the memory module includes the power of each component, the spacing between the memory modules, air flow velocity and temperature, as well as the presence of any thermal solution (e.g. heat spreader).

DIMM memory is typically downstream from the processor, hard disks and fans, and therefore has a local higher ambient temperature. In typical server system layouts, cool air will flow from one end of the DIMM to the other, with the hottest DRAM component usually being the last one on the same side as the Register. However this conclusion is not consistent across all DIMM formats. For example, the fully buffered DIMM's hottest DRAM is near the center of the DIMM card, next to the Advanced Memory Buffer.

Memory thermals are important because good memory thermals improve system performance in addition to requiring less power. The thermal characteristics of memory subsystems are important because when memory subsystems operate at lower temperature, system performance improves and overall system power consumption is less. The memory thermals are characterized as a function of fan speed and preheat to the DIMMs. The required cooling capability in watts per DIMM varies depending upon whether the DIMM has a Full DIMM Heat Spreader (FDHS) or not and whether the DIMM is in double refresh or not.

A DIMM under double refresh has a case temperature specification of 95°C rather than 85°C, thereby enabling a higher overall safe system temperature at the expense of a small power loss and slightly increased power consumption. The impact of double refresh (85°C versus 95°C) is substantial improving cooling capability by approximately two to three watts resulting in a significant improvement in memory bandwidth capability.

Throttling Memory to Reduce Power Consumption

Intel processor based servers include automatic memory throttling features to prevent memory from overheating without the processor or memory using additional power. There are two different memory throttling mechanisms that are supported by Intel chipsets: closed loop thermal throttling (CLTT), and open loop throughput throttling (OLTT).

Closed loop thermal throttling is a temperature-based throttling feature. If the temperature of the installed FB-DIMMs approaches their thermal limit, the system BIOS will initiate memory throttling to manage memory performance by limiting bandwidth to the FB-DIMMs, therefore capping the power consumption and preventing the FB-DIMMs from overheating. By default, the BIOS will configure the system to support CLTT if it detects that there are functional advanced memory buffer (AMB) thermal sensors present on all installed FB-DIMMs. In CLTT mode, the system fans run slower to meet the acoustic limits for the given platform but will also allow the fans to ramp up as needed to maintain the parts within temperature specifications under high stress levels.

Open loop throughput throttling (OLTT) is based on a hardware bandwidth count and works by preventing the bandwidth from exceeding the throttling settings programmed into the MCH registers. The system BIOS will automatically select OLTT as the memory throttling mechanism if it detects that one or more installed DIMMs does not have a functional AMB thermal sensor. Once the system BIOS enables OLTT, it utilizes a memory reference code (MRC) throttling algorithm to maximize memory bandwidth for a given configuration. The MRC code relies on serial presence detect (SPD) data read from the installed DIMMs as well as system level data as set through the FRUSDR Utility.

While memory throttling is good in that it prevents memory failures without consuming additional power, it has limitations in that it can negatively impact system performance. Program execution can be affected when the memory is shut down or when the memory bandwidth is limited by CLTT or OLTT.

Power Supplies

Power supplies transform AC power into DC for use by server circuitry and power transformation loses some energy. The efficiency of a power supply depends on its load. The most efficient loads are within the range of 50–75 percent utilization. Power supply efficiency drops dramatically below a 50 percent load, and it does not improve significantly with loads higher than 75 percent.

Power supplies are typically profiled for efficiency at a very high load factor, typically 80 to 90 percent. However, most data centers have a typical loads of 10 to 15 percent. Thus power supply efficiency is often poor. Since most servers today run with 20-40 percent efficient power supplies, they waste the majority of the electricity that passes through them. As a result, today's power supplies consume at least 2 percent of all U.S. electricity production. More efficient power supply designs could cut that usage in half, saving nearly \$3 billion.

A high efficiency power supply can significantly reduce overall system power consumption. For example for a 400W system load, a 60 percent efficient supply consumes 560W at the wall vs. 460W with an 85 percent efficient power supply. Potential power savings from the change to a more efficient power supply = 100W.

In addition to the main power supply, servers utilize secondary power supplies that also can waste some power. These secondary smaller power supplies are distributed across the motherboard and are located close to the circuits they power. Secondary power supplies used in servers include point-of-load (POL) converters, voltage regulator modules (VRM) and voltage regulator down (VRD).

The output voltage from a VRM or VRD is programmed by the server processor using a voltage identification code (VID). Other secondary power supplies, such as POL converters, do not have this feature. VRM and VRD voltage and power requirements will vary according to the needs of different server systems. In many servers, approximately 85 percent of the motherboard power is consumed by the VRM/VRD - exclusively for the server's processor.

To minimize power consumption with power supplies and the secondary voltage regulators, a server should run workloads to optimize the power supply efficiency. Intel multi-core processors work with the VRM/VRDs so each core can operate in the most efficient way.

Storage Systems and Power Consumption

A basic server with two or four hard disk drives (HDDs) will consume between 24W and 48W for storage. By themselves, a few disks do not consume that much power. But external storage systems in large enterprises have thousands of disks

that consume significant amounts of power in the data center. Small businesses typically purchase servers with direct-attached storage, where the server contains many HDDs. Increasingly, small businesses also purchase networked storage systems shared by client and server systems.

Fewer storage devices consume less energy. Better utilization is the key. Poor storage management practices can consume significant amounts of power. The most common wasteful storage practice is operating disks that manage low-activity data spinning 24 hours per day. Underutilization of data access (and thus the data's value) increases power and cooling expenses when compared to better managed storage solutions that use energy only when data is accessed or written.

Storage utilization figures differ by operating system and type of storage device. On typical server systems, the average usage level for a hard disk is about 40 percent. New disk drive capacity is increasing much faster than drive performance. As a result of this imbalance, storage administrators typically use the redundant array of inexpensive disks (RAID) architecture and striping techniques to increase performance and reliability, but at the price of increasing the number of rotating drives. As utilization levels drop, more devices are needed, increasing total disk costs and energy expense.

Rotating drives consume energy and generate heat. Hard disk drives, like other computer components, are sensitive to overheating. Manufacturers measure off a minimal range of operating temperatures - from +5 to +55°C as a rule (occasionally from 0 to +60°C), which is a smaller range than for processors, video cards, or chipsets. The reliability and durability of HDDs depend on operating temperatures. Increasing HDD temperature by 5°C has the same effect on reliability as switching from 10 percent to 100 percent HDD workload. Each Centigrade degree drop of HDD temperature is equivalent to a 10 percent increase of HDD service life.

The heat rate dissipated by a HDD is the product of its current and voltage in various states. The efficiency of HDD small motors can be less than 50 percent. Power consumption of hard drives is usually measured in states of Idle, SATA or SCSI Bus Transfer, Read, Write, Seek, Quiet Seek (additionally, if supported), and Start. Average power consumption of a HDD can be calculated by measuring the power consumption of a HDD both during typical user operations and during intensive (constant) operations.

For every usage model, the percentage idle versus active of a HDD depends on disk capacity, applications in use, and related workloads. Average power consumption is estimated with the formula noted below, but it should be noted that your mileage may vary for actual power use.

Average hard disk power consumption for average operations, such as office work, $P_{Average}$ can be estimated by the following formula:

$$P_{Average} = \frac{P_{Idle} \times 75\% + P_{Write} \times 5\% + P_{Read} \times 20\%}{100\%}$$

where the states represent the power consumption of a drive from the voltage sources and the percentages represent the common percentage of the HDD state duration. This formula is based on the assumption that read/write HDD operations make up 10 percent of the total time for the average office usage.

Average power consumption during intensive hard disk operations such as defragmenting disks, scanning the surface, copying files ($P_{Constant}$) can be calculated by the following formula:

$$P_{Constant} = \frac{(P_{Write} + P_{Seek} + P_{read}) \times 4}{5}$$

This formula is based on the assumption that read/write HDD operations in intensive operations is over 50 percent of the time.

The most efficient hard drives will consume on average 5 - 6W in idle state. Average SATA interface hard drives consume between 7 and 10W in idle. Today's SATA hard drives typically consume between 10 and 15W during active modes. And, as with other computer components, the efficiency of the latest generations is much better than those of previous generations.

Heat dissipation requirements have relaxed because power usage of hard disks has been steadily going down in recent years. Newer serial interfaces (e.g., SATA II and SAS) do increase power usage and heat dissipation a bit, but overall the heat dissipation trend is moving down. And, Quiet Seek mode (i.e., slowing down the drive so the noise of rotating platters is equal or less than 128 decibels) can sometimes reduce heat dissipation of a hard disk much lower than it is increased by using a newer serial interface.

Summary

Electricity used by servers has doubled between 2000 and 2005 due to the increase in the number of servers installed and the required cooling equipment and infrastructure. Power use varies by server type, the configuration within each server, and the workload being run. All server components generate heat as they function. An increase of the local ambient temperature inside a server can cause reliability problems with the circuitry. Additional power is needed to keep systems

and their components within a safe operating temperature range. As data centers move towards increased server density, the power and heat generated by servers will increase.

On a basic server system today, server processors consume the most power, followed by memory, then disks or PCI slots, the motherboard, and lastly the fan and networking interconnects. The recent shift to multi-core processors helps address energy consumption by the CPU and adds power management features that throttle processor power.

But today's applications are far more processor-intensive, which has triggered a trend toward high-density packaging and increased memory. This trend leads to memory as the largest power consumer in servers in the years to come. Memory cooling thus emerges as the primary thermal challenge.

Power supplies waste energy by converting AC power into DC. Most servers run today with 20 - 40 percent efficient power supplies that waste over half the electrical power passed through to them. At the typical server level, potential power savings from the change to a 15 percent more efficient power supply can be as great as 100W or more.

Storage power is minimal for an individual hard disk drive, but when servers utilize several disks and interoperate with RAID arrays and networked storage systems, storage power consumption is significant. Greater and cheaper disk drive capacity without performance improvements is leading to the trend of redundant disks and striping for performance and reliability. Good storage management techniques can offset this trend.

Power demands are trending upward, but newer server components are being manufactured to run more efficiently. The latest Intel Xeon processors, power supplies, memory and even hard disk drives all use less power, include power management features, and create less heat. With newer servers, data centers can evaluate servers by their performance per watt and focus on business optimization instead of TCO costs.

For more information about server power consumption, please refer to the book *Energy Efficiency for Information Technology* by Lauri Minas and Brad Ellison.

About the Authors

Lauri Minas is the Senior Strategic Server Planner in the Server Architecture and Planning Group within Intel's Server Products Group. Lauri has managed Intel's Server Industry Marketing Group, setting the strategic direction of server industry

efforts for Intel server technologies across Intel divisions. She is a five-time recipient of the Intel Achievement Award. Lauri earned both her bachelor's degree and master's degree in Business from Arizona State University.

Brad Ellison is the Data Center Architect in the IT Ops Data Center Services team within Intel's IT organization. Ellison has sat on the Data Center Institute's Board of Directors and is a charter member of the Infrastructure Executive's Council's Data Center Operations Council and the ICEX Knowledge Exchange Program Data Center Excellence Practice. Ellison received a B.S. from the University of North Dakota in 1981 and an M.S. from Oregon State University in 1985

Copyright © 2009 Intel Corporation. All rights reserved.

This article is based on material found in book *Energy Efficiency for Information Technology* by Lauri Minas and Brad Ellison. Visit the Intel Press web site to learn more about this book: http://www.intel.com/intelpress/sum_rpc.htm

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4744. Requests to the Publisher for permission should be addressed to the Publisher, Intel Press, Intel Corporation, 2111 NE 25 Avenue, JF3-330, Hillsboro, OR 97124-5961. E-mail: intelpress@intel.com .