Server Memory Solutions for the Impending Data Center Power Crisis

Rahul Advani
Micron Technology, Inc.
May 5, 2008

Executive Summary

Until very recently, data center managers were only concerned about performance and acquisition costs when they were considering buying new servers or expanding their facilities.[1] But soaring energy costs and system power consumption have forced data center managers to analyze all aspects of their electrical and thermal expenditures—including server memory modules—as these managers seek to save money, become more efficient, and act to conserve resources for future generations.

Recognizing this challenge, Micron's research and development engineers discovered a way to reduce power consumption in memory modules without affecting performance. These low-voltage DDR2 and DDR3 SDRAM solutions significantly reduce power consumption (perhaps 15–20%) and generate less heat. Put another way, standard DDR2 and DDR3 SDRAM devices generally operate at 1.8V and 1.5V, respectively, but by making subtle but complicated changes in semiconductor architecture, Micron has been able to adjust these devices to operate at reduced voltages of 1.5V and 1.35V, respectively. These lower operating voltages translate into both reduced power consumption and less heat.

Data Center Quick Facts

- From 2000 to 2005, total worldwide server energy consumption nearly doubled, accounting for 0.8% of the world's total power usage in 2005 when auxiliary infrastructure and cooling are considered.[3]
- In 2006, 1.5% of all the power consumed in the U.S. went to data centers.[4]
- Unless action is taken, 90% of data centers will experience a power-related failure in the next five years.[5]
- A 2005 survey of IT professionals showed that power and cooling were the primary issues in data center management.[6]

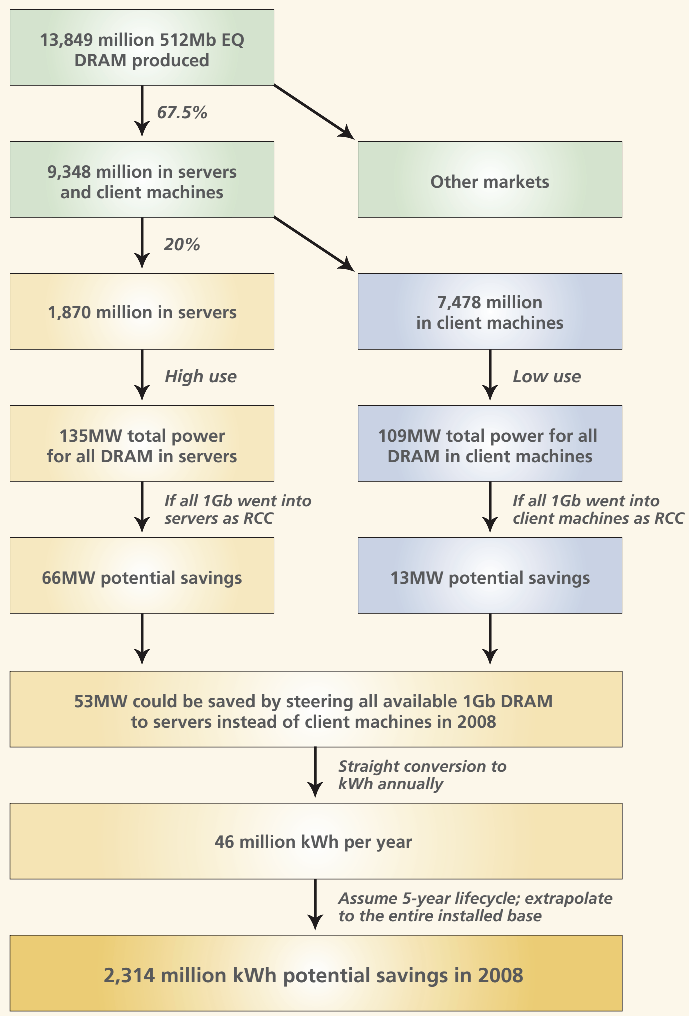


Figure 1: Estimated savings from implementing 1.5V DDR2 SDRAM based on annual DRAM manufacturing and market usage[2]

- Using more efficient DRAM in server modules could save data centers more than $300 million per year in power costs.[7]

Purpose

This paper is an update and expansion of Michael Sporer's white paper, "The Power Demands of Data Centers Require Memory Innovations." Like Sporer's work, this paper looks at what is potentially a power and heat crisis for data centers worldwide. While there is no single panacea to address a data center's challenges, there are several small steps that can be taken and that will add up to huge savings.

The Data Center Defined

This paper defines a data center as a room(s), floor(s), entire buildings, or groups of buildings that house computer, storage, and networking equipment and the infrastructure needed to support that equipment. Data centers generally act as the backbone for the Internet and corporate networks, serving up Web pages, stream media, or providing voice over Internet protocol communications (VOIP). Data centers can and do run simulations of all kinds and support a plethora of traditional and private computing functions like banking, e-commerce, music downloads, and more.[8]

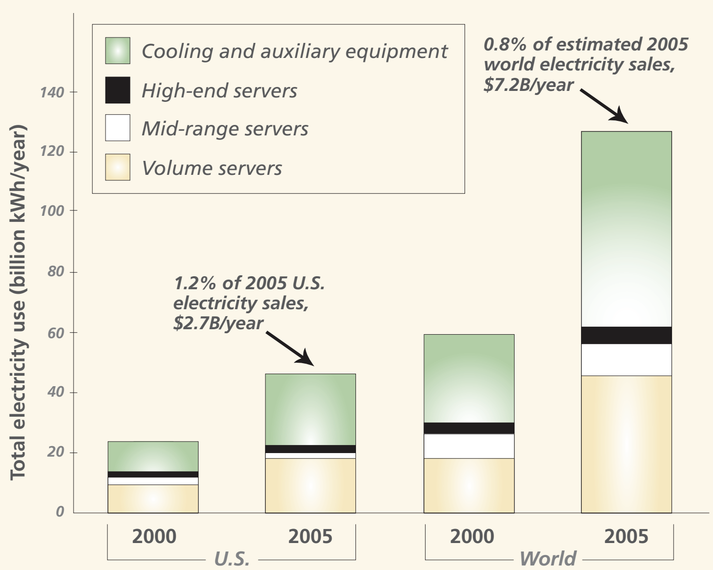
The Coming Data Center Power Crisis

"Over the next five years power failures and limits on power availability will halt data center operations at more than 90% of all companies."[9]

Instead of paper letters, people communicate with email. Instead of ledgers, banks use massive financial networks. Everything from television to how consumers get music has been shifting to digital-and-network-based systems, effectively putting more and more pressure on data centers.[10]

Consider that Internet usage increases about 10% each year, representing a 13–20% CAGR.[11] This increased use has a huge impact on just how much power data centers consume. In fact, from 2000 to 2005 data center power consumption doubled to nearly 45 billion kWh in the U.S. alone.[12] The total electric bill for data centers in 2005 was about $2.7 billion for the U.S. and about $7.2 billion for the entire world.[13]

This massive boost in data center demand and therefore data center power consumption cannot be overlooked. It is a crossroad that—if ignored—could lead to widespread system failures, business interruptions, or lost opportunities.



Sources: IDC data for installed base, shipments, and most popular models, and manufacturer data on power use for individual server models. Total expenditures assume U.S. industrial electriciy prices (2006 dollars)

Figure 2: Total electricity use for servers in the U.S. and the world in 2000 and 2005, including the associated cooling and auxiliary equipment.[14]

Here are some specific and public examples of the sorts of power problems that data centers will face.

"The University at Buffalo installed a $2.3 million Dell Inc. supercomputer [in the summer of 2005] hoping to bolster its image as a research institution. Instead, the big machine came to symbolize an increasingly vexing problem for data centers worldwide. Once the machine was delivered, university officials discovered they had only enough electrical power to switch on two-thirds of the system."[15]

"New, more energy-efficient machines can't come fast enough for Denis Weber, executive director for information technology at Verizon Wireless. The company, a joint venture of Verizon Communications Inc. and Vodafone PLC, was forced to upgrade a data center in Ohio to bring in two megawatts of power, a nearly sevenfold increase from 1998, he says. The monthly power bill for the center rose to $40,000 over the same period from about $10,000."[16]

"Consider the experience of Pomona Valley Medical Center. The hospital east of Los Angeles quickly grew to 70 servers from 30, says Kent Hoyos, its chief information officer. The heat they generated overwhelmed the computer room's air-conditioning system and a backup unit that was pressed into service. With temperatures in the room averaging 92 degrees, machines began behaving erratically, Mr. Hoyos says. In late 2003, an air-conditioning unit broke down, sending the temperature over 100 degrees. The event caused a temporary shutdown of systems serving the hospital's laboratory, $40,000 in damage to servers and hard drives, and prompted a $500,000 retrofitting of the cooling system."[17]

Projections for Future Consumption

Generally, data center power consumption can be divided into three areas: data center server load consumption; server cooling systems; and other data center infrastructure, including lighting, office space, and electrical room cooling. In the situations mentioned above we can see examples of each of these types of challenges, which, if ignored, will have potentially catastrophic consequences as more and more demand is placed on data centers.

Because of the IT industry's propensity to innovate, the would-be data center crisis is not as bad as it could be. For example, when the U.S. Environmental Protection Agency's (EPA) offered its comprehensive Report to

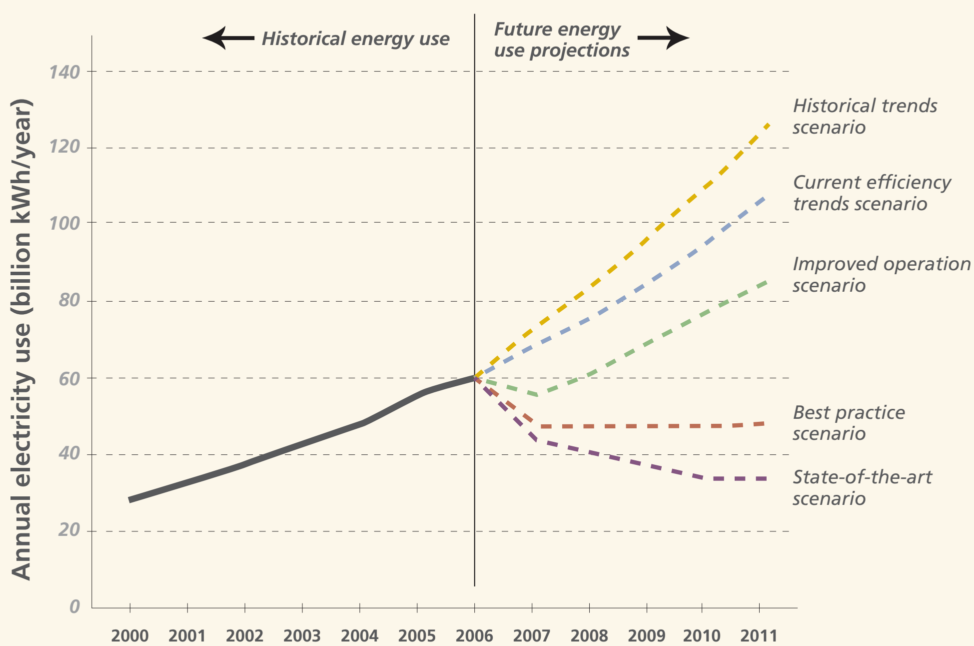


Figure 3: Comparison of projected electricity use at data centers for all scenarios[22]

Scenario	*Electricity consumption savings (billion kWh)*	*Electricity cost savings ($billion 2005)*	*Carbon dioxide emissions avoided ($MMTCO_2$)*
Improved operation	23	1.6	15
Best practice	60	4.1	38
State-of-the-art	74	5.1	47

Table 1: Annual savings in 2011 by EPA scenario when compared to current efficiency trends[23]

Congress on Server and Data Center Energy Efficiency in 2007, it concluded that there were five scenarios for projecting future data center power consumption. The first of these was an historical perspective that simply used data from 2000 to 2005 to estimate what data center power consumption would look like in 2011. But the agency's research indicated that data centers had become more energy efficient since 2005, necessitating a "current" scenario which took into account the power-saving technologies and policies that were already being implemented.[18] The EPA also named three other scenarios for future data center power consumption. The improved operation case includes improvements and innovations that exceed current trends in terms of energy efficiency, but are available and could easily be implemented.[19] The best practices case "represents the efficiency gains that can be obtained through the more widespread adoption of the practices and technologies used in the most energy-efficient facilities in operation today."[20] And finally, the state-of-the-art scenario includes using the most advanced available technologies and practices to maximize efficiency in data centers.[21]

It becomes clear that in order to achieve what the EPA described as the "state-of-the-art" scenario every possible means must be employed to achieve the maximum energy efficiency possible,[24] including selecting the most power-frugal memory.

Memory's Role in Servers and Energy Use

Whether in use or not, memory always draws power, so if a server is on and idle, the system memory is still refreshing every few nanoseconds.[25] In DRAM, a transistor and a capacitor make up a memory cell. Each memory cell represents a single bit of data stored as either a '0' or a '1.' In order to keep that data valid, each memory cell must be refreshed (given a fresh jolt of power). DRAM uses more power when it is being written to or read from than when it is idle, but nonetheless represents a steady demand for electricity.

As this paper is an extension of Sporer's earlier work, it is reasonable to quote that earlier work at length.

Another factor driving memory content growth and server power consumption is the adoption of virtualization technologies. A server running a virtualized environment is able to achieve a higher utilization which, in turn, increases the total power consumption of the server. Once again, the importance of energy efficient component selection increases. By analogy, a car will burn very little fuel if it isn't driven. Virtualization, or anything that increases server operation, is like adding a new driver to the mix. Now the car gets driven more and energy efficiency becomes a greater concern.

Traditionally, the CPU has been the component that consumes the most power in the system. Improvements in CPU power consumption now place a greater scrutiny on the other components. Where memory once played a distant second to CPUs in the ranking of system power consumption, now, in some cases, it exceeds the power consumption of the CPU.[26]

Assuming that memory represents about 20% of the total power budget in a server, Sporer estimated that using more efficient 1.5V DDR2 memory modules would save

Component	Peak Power (Watts)
CPU	80
Memory	36
Disks	12
Peripheral slots	50
Motherboard	25
Fan	10
PSU losses	38
Total	251

Table 2: Component peak power consumption for a typical server[27]

24% of the total server demand for power and dramatically reduce the heat the system generates.[29]

Next-Generation, Low-Voltage Memory Solutions

In 2007, Micron introduced a family of energy-efficient Aspen Memory® products to address rising concerns about server power consumption. The first devices available in the Aspen Memory family were based on 1Gb DDR2 SDRAM that operates at 1.5V rather than the standard 1.8V. These modules were able to reduce a server's power consumption by as much as 60%—20% due to the reduced-voltage components and an additional 40% due to their reduced chip count design. Egenera, a leader in server development and server virtualization, has already begun using Aspen Memory DDR2 in its most energy-efficient servers.

In 2008, Micron expanded their Aspen Memory family to include low-voltage (1.35V), DDR3 SDRAM modules and innovative 2Gb-based DDR2 modules. The DDR3 devices offer the most efficient, best performing memory technology a server could use, while the 2Gb-based DDR2 modules will dramatically reduce the number of components needed in a memory module, significantly lowering system heat.

Together these innovations are helping data centers achieve the state-of-the-art power savings that will effectively reduce the requirements for data center power upgrades and send total data center power consumption below 2005 levels without compromising performance.

Conclusion

There is a coming data center power crisis, but leading companies and the U.S. government are taking data center power consumption very seriously, even proposing scenarios where data center power use could fall below historical levels. Micron is helping to defuse this potential power crisis by leading the world in the development of energy-efficient server memory.

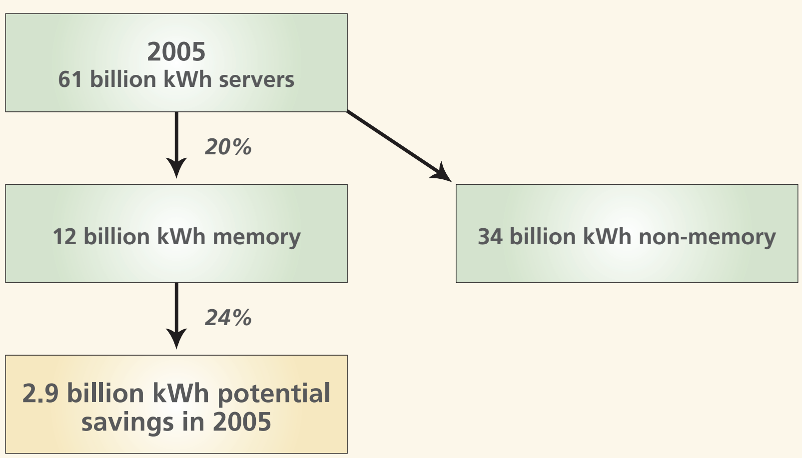


Figure 4: Estimated memory power consumption and potential savings[28]

[1] *"Curtailing the Power and Cooling Epidemic," Egenera, Inc., Marlboro, MA, 2007, page 2.*
[2] *Michael Sporer, "The Power Demands of Data Centers Require Memory Innovations," Micron Technology, Inc., 2007, page 6*
[3] *Jonathan G. Koomey, Ph. D., "Estimating Total Power Consumption by Servers in the U.S. and the world," Lawrence Berkeley National Laboratory, Berkeley, Calif. February 15, 2007, page 17*
[4] *EPA, page 18*
[5] *AFCOM*
[6] *EPA, page 18*
[7] *Sporer, page 1*
[8] *Sporer, page 2 and Micron research*
[9] *"Five Bold Predictions for the Data Center Industry that Will Change Your Future," AFCOM Data Center Institute, March 2007, http://www.afcom.com/files/PDF/DCI_Keynote_Final.pdf, (accessed on April 25, 2008)*
[10] *Koomey, pp. i and ii*
[11] *EPA, page 27*
[12] *Koomey, page 6*
[13] *Koomey, page i*
[14] *Koomey, page 17*
[15] *Don Clark, "Power-Hungry Computers Put Data Centers in Bind," Wall Street Journal, New York, November 14, 2005, page A1*
[16] *Ibid.*
[17] *Ibid.*
[18] *EPA, pp. 7-10*
[19] *Ibid.*
[20] *Ibid.*
[21] *Ibid.*
[22] *Ibid.*
[23] *Ibid.*
[24] *Ibid.*
[25] *"Mark Blackburn, "Five Ways to Reduce Data Center Server Power Consumption," The Green Grid, Beaverton, Ore., 2008, page 4*
[26] *Ibid.*
[27] *EPA, page 19*
[28] *Sporer, page 5*
[29] *Sporer, page 6*

micron.com