WHITE PAPER

The Invisible Crisis in the Data Center: The Economic Meltdown of Moore's Law

By Kenneth G. Brill

A continuing gap, at the power plug, between a faster rate of computational increase and a slower rate of energy efficiency improvement will ultimately cause the economic meltdown of Moore's Law.

—Kenneth G. Brill, 2007

Uptime Institute™

This Paper:

1. Explores the repercussions of Moore's Law from the perspective of CIOs and CFOs on the net economic productivity of enterprise IT operating as a whole system, including data center site infrastructure (power and cooling)
2. Explains why facilities' costs (site OpEx and CapEx) are growing from the historic one to three percent of the total IT budget to five to 15 percent with no upper boundary in sight
3. Considers how this major economic change could occur quickly and invisibly
4. Outlines necessary changes in IT governance and computer room management to avert a meltdown (both figuratively and literally), postulates a new technology roadmap for chip, server, and storage manufacturers that will permit energy efficiency to rise at a faster rate than computational performance, poses the question of whether IT budgets will be increased to compensate for growing data center site and energy costs or whether increasing site TCO will force IT abandon otherwise valid initiatives or significantly reduce hardware spending

Background

Our use of the term *meltdown* is for dramatic effect. For some enterprise IT organizations, it will be hyperbolic, for others it will have the ring of absolute truth. Our opinion, based on available data, is that a largely unseen cost crisis is looming within data centers, fueled by a growing divergence between rapid server computing performance gains and a far slower improvement in energy efficiency. The resulting increase in server power consumption now burdens data center operators with increasing electricity and site infrastructure costs for every server they deploy.

In this white paper we examine the economic productivity of large-scale server computing and the enterprise data center.

In the past, rising IT power consumption trends were simply background noise that received (and warranted) scant senior user executive attention. However, as processor manufacturers such as Intel, AMD, IBM, and others continue to deliver on Moore's Law—doubling the number of transistors on a piece of silicon every 18 months (even faster than the 24 months that Intel co-founder Gordon Moore originally predicted in 1965)—the resulting power density increase within chips is causing the temperatures inside and around those chips to rise dramatically.

The chip-level density problem is further compounded by server hardware manufacturers packing increasing numbers of ever–more power-hungry chips into the same or smaller footprints. Virtually everyone involved in large-scale IT computing is now aware of the resulting temperature and cooling problems data centers are experiencing, however, very few understand that the meltdown is fundamentally a high-stakes economic problem, rather than an engineering problem (which can be solved with time and money).

Meltdown in this context means that Moore's Law is affected when the problem is examined at the whole-systems level of the server farm and enterprise data center (composed of mainframes, servers, disk and tape storage, communications and networking equipment), rather than at the level of an individual computing device. The widening delta between the faster rate of increase of server computational performance and the slower rate of increase in energy efficiency growth results in a dramatic increase in user engineering and cost burdens of providing power and cooling.

Root of the Problem

In mathematical terms, server compute performance has been increasing by a factor of three every two years, however, energy efficiency is only doubling in the same period. This means computational performance increased by a factor of 27[1] between 2000 and 2006. Energy efficiency has gone up as well, but by only a factor of eight[2] during the same period.

This means that while power consumption per computational unit has dropped dramatically in this six-year period (by 88 percent), the at-the-plug power consumption has still risen by a factor of 3.4[3].

A constant rate of IT hardware spending results in increasing hardware power consumption at-the-plug, which, in turn, results in rapidly escalating site electric utility costs as part of the enterprise IT operating expense. This trend is largely invisible until an economic crisis occurs that attracts the attention of C-suite executives. This typically happens when an enterprise data center runs out of power and/or cooling capacity, and an unplanned (or sub-optimally planned) site capital investment is required to increase capacity. For large-scale enterprise data centers, these CapEx investments can now be in hundred million dollar increments.

One would think that with financial and business performance risks of this magnitude, data center owners would have a high-level of awareness and management preparedness. However, a recent survey of 100 data center operators by the Aperture Research Institute suggests something quite different for many organizations.

Almost 40 percent of respondents reported that they had run out of space, power, or cooling capacity without having sufficient notice. This can result in delaying the provisioning of new initiatives or paying more for additional resources to host those systems.

—Aperture Research Institute. 2006. "Organizations Struggle with Data Center Capacity Management."

The Aperture survey also revealed that 11 percent of the respondents admit that their infrastructure capacity planning procedures are *"poor"*; 31 percent say they are only "*fair*", with less than 20 percent giving their management processes an *"above average"* ranking.

1 *There are three two year cycles between 2000 and 2006. Three times three times three equals 27.*

2 *Two times two times two equals eight.*

3 *Twenty-seven divded by eight equals 3.375 or 3.4*

Table 1 shows the economic consequences of spending $1M annually for 1U servers in different years. The number of compute performance units goes up by a factor of 27 between 2000 and 2006. The *Institute* estimates that during the same period, power consumption goes up by factor of 3.9. (This is greater than the 3.4 discussed earlier because server prices are dropping by around three percent per year and $1M of hardware spending in 2006 buys more servers each consuming more power.) While power per performance unit drops significantly, total power consumption still goes up dramatically.

Table 1 also reports initial capital costs for site infrastructure (data center power, cooling, and physical building), three-year electricity costs, and three-year site TCO costs. (We choose three years because this is the functional life of most servers). In 2000, these costs could easily be overlooked by senior executives as insignificant, however, in a period of just 12 years (2000 to 2012), the consequence of computational performance increasing at 3X and energy efficiency increasing more slowly at 2X every two years is a 15X escalation in the three-year site TCO associated with a constant $1M in server spending. (Is your organization spending more or less money during each server technology refresh cycle? Many companies are spending more, not less making the rising burden of site TCO even more serious.) If current trends continue, by 2012, $1M of server spending will invisibly commit an organization to $6.54M more in three-year site infrastructure TCO than the same $1M of server purchasing required in 2000. As a result, site infrastructure cost as a percentage of IT's total budget is rapidly escalating with no upper limit in sight. Of particular interest to C-suite executives is that invisibly rising infrastructure costs are necessarily draining money from other IT initiatives, unless the IT budget is also increased. Increasing IT's share of total enterprise revenue does not seem viable for most businesses.

The site OpEx and CapEx numbers for *Table 1* are taken from *Institute* benchmarking within its Site Uptime Network® and from other *Institute* white papers (primarily, *Dollars per kW plus Dollars per Square Foot Are a Better Data Center Cost Model than Dollars per Square Foot Alone*). Linking site costs over the life cycle of the server helps illustrate the economic consequences of current trends.

Table 1—Site TCO with a Constant $1,000,000 Rate of 1U Server Spending

Year	Compute Units per $1.0M (Year 2000 = 1)	Server Spend	Server kW	Site CapEx	3-Year Site Site Electric	3-Year Site TCO	3-Year Site TCO / Server Spend
2000	1	$1.0M	32 kW	$0.770 M	$0.130M	$0.460M	46%
2003	5	1.0	63	1.500	0.260	0.910	91
2006	27	1.0	125	3.000	0.500	1.810	181
2009	140	1.0	242	5.800	0.980	3.510	351
2012	729	1.0	482	11.600	1.950	7.000	699

Table 1 Notes and Assumptions

1. In 2006, a 1U server has a street price of $4K and has an actual power consumption of 500 Watts *(Belady, Christian. 2007. "In the Data Center, Power and Cooling Costs More Than the IT Equipment it Supports." Electronics Cooling. Vol. 23, No. 1, February 2007).* This indicates that $1 M in server spending buys 250 servers, which consume a total of 125 kW.
2. From 2000–2012, server performance increases by 3X every 2 years, while performance per Watt only doubles (Belady, 2007). This means that power consumption per server goes up by a factor of 1.2 each year. During this period, the cost per server falls by 3 percent per year.
3. Site CapEx cost is $24K per kW (assuming 100W per ft2 of gross computer room) of power delivered to IT hardware, typical for a Tier IV data center. This is the up-front (not amortized) capital expenditure for construction (not including significant multi-million dollar costs like land, architect and engineering fees, permits, IT costs like migration, tenant fit-out, and network extension, and interest during construction.) The $24K per kW CapEx figure is based on 2005 benchmarking which has significantly escalated over the last two years.
4. Server life is three years, while site infrastructure life is 15 years.
5. The electricity cost is based on a constant rate of $0.07/kWh including both energy and kW-based demand charges. This rate is reflective of Site Uptime Network members and the national average. Many regions pay significantly more ($0.14 for New York City). Site infrastructure has a Site Infrastructure Energy Efficiency Ratio (SI-EER) of 2.2, indicating that 2.2 kW are consumed at the utility meter for every kW delivered to IT server load.
6. Three-year site TCO includes OpEx and amortized CapEx. Costs used in Table 1 significantly under estimate what many companies will actually incur. Site OpEx is $3,237 per year, per kW of power delivered to the IT hardware and is based on Site Uptime Network member data for three-shift coverage with a minimum of two technicians per shift for sites located in non-union, low-cost areas. This number includes electricity, the site technicians required to operate and maintain power and cooling equipment, and external maintenance and repair costs for site infrastructure equipment and systems. Site OpEx does not include property taxes, interest, insurance, and CapEx depreciation. TCO also includes amortized CapEx of 20 percent of the initial CapEx (three years out of the 15-year life).
7. Calculations assume a just-right match of IT hardware kW load and site capacity. For the typical 50 percent excess site capacity, increase the 3-year site TCO column by a factor of approximately 1.3.
8. Many companies will incur costs significantly greater than those shown in Table 1.

When looking at *Table 1*, note that the problem does not go away if slightly different assumptions or numbers are used. For instance, if we assume that servers in 2006 use only 320 Watts and still cost $4K, then site electricity costs will exceed server purchase price in 2012 rather than 2009. Similarly, assuming that a 2006 server uses 500 Watts but that server power use only grows by a factor of 1.1X per year (instead of 1.2X per year) also delays the date at which site electricity costs will exceed server purchase price to 2012. So more optimistic assumptions only delay the problem a few years rather than making it go away. Eventually, this is something that the uninterruptible uptime industry will be forced to resolve. (And, of course, if the forecasts we derive from Belady are low, the meltdown will loom even sooner than we think.)

Will Virtualization Save Us?

It may be tempting to think that virtualization[4] alone can relieve the burden of rising site infrastructure expenses, but unfortunately this is not the case. Virtualizing four or 10 servers onto a single box will indeed cut power consumption and free up precious data center capacity. For data centers nearing their limits, virtualization can play a key role in delaying the time at which an expansion or new facility must be built. But that's only half the story.

In many ways, the "economic meltdown" is defined by the point at which the costs of electricity and site infrastructure TCO greatly outpace the cost of the server itself. This can happen regardless of whether a single box is running one application or virtualized to handle multiple tasks. When electricity and infrastructure costs greatly exceed server cost, any IT deployment decisions based on server cost alone will result in a wildly inaccurate perception of the true total cost. Even when virtualization frees up wasted site capacity for additional servers without spending new money on site infrastructure, the opportunity cost of wastefully deploying this capacity is the same; data center managers will find themselves building expensive new capacity sooner than they need to.

Furthermore, virtualization is a one-time benefit. After consolidating servers so that they are all running at full capacity, and planning future deployments so that newly-purchased servers will also be fully utilized, data center operators are still faced with the reality that each year's generation of servers draws more power than the last. After virtualization has taken some of the slack out of underemployed IT hardware, the trend in power growth will resume.

To be fair, it's possible that virtualization may allow each new server to be so productive that it's worthwhile to divert a greater fraction of the IT budget to pay the increased site infrastructure and electricity cost, but a business can't make that decision without knowing the true total cost.

By 2009, the Three-Year Cost of Electricity per Server Will Exceed the Purchase Cost of the Server

The numbers chosen for site costs are very conservative and significantly underestimate what many companies are already paying. For example, electricity was fixed at $0.07 per kWh. Many data centers in the northeastern US will experience rates double this or more, and this will continue to escalate as 2012 nears. Just using seven cents as a fixed rate, sometime around 2009, the three-year cost of the subsequent electricity will exceed the street price (as opposed to *manufacturer's suggested retail price* which will be significantly higher) of the server itself. Perhaps more immediately obvious to Facilities and IT personnel is the fact that power-hungry IT hardware quickly gobbles up expensive UPS, engine generator, and cooling capacity. Once this capacity is exhausted, additional productive IT equipment cannot be deployed until the capital (and time) is spent to expand an existing data center or build a new facility and move IT. Saving on the electricity bill is important, but saving precious site capacity is often the primary benefit of increased energy efficiency.

The cost categories for the three-year TCO used in *Table 1* are electric utility (28 percent), site technicians and operators to run power and cooling equipment (18 percent), other site equipment and systems maintenance costs (21 percent), and amortized CapEx (33 percent). Costs for physical security, server licenses, applications, network, communications, help desk, break/fix, system administrators, and data center management are not included in these site numbers.

The capital expenditure for a new data center flows out at the time of construction. Different companies will have widely differing policies for how they account for such expenditures. As used in this paper, amortized CapEx can be simply thought of as three years of straight-line 15-year depreciation. The 15-year life was selected because most site infrastructure systems need refurbishments after five IT hardware obsolescence/refresh cycles or 15 years. The method for calculating total CapEx investment uses 2005 construction costs (which have since sharply increased), Tier IV site design topology (see *Tier Classifications Define Site Infrastructure Performance* white paper), and a perfect match of hardware load and site capacity. In reality, the CapEx costs shown significantly underestimate the actual costs incurred by many companies.

Not all companies require the fault tolerance and concurrent maintainability of a Tier IV data center. However, with amortized CapEx being only 33 percent of 3 year-site TCO, even a Tier II data center which only requires 50 percent of a Tier IV facility's CapEx will still experience similar escalating site TCO costs.

4 *The* Institute *uses virtualization here as an example, but this same logic applies to application rationalization, "right sizing" of server hardware, virtualization of storage, and a myriad of other ways of better utilizing existing compute capacity.*

Embedded Watts per $1,000 of Hardware Spending

A simple and effective way to project future site TCO costs is to develop and maintain a simple chart of hardware power consumption at-the-plug divided by the actual price paid for the hardware in thousands of dollars. This is in addition to the traditional IT metrics used for price/performance.

The following are three very important precautions:

- Do not use the power consumption figures shown on the IT hardware equipment nameplate. These figures will significantly overstate actual power consumption (by up to 50 percent).
- Do not use manufacturers' suggested list prices. Significant discounts are typically available to major users. Use street or actual prices paid.
- Develop separate spending charts for different types of IT equipment. Within servers, this means blades, 1U volume servers, midrange servers, and high-end servers. For disk storage, look at spindle speed. Communications/networking equipment must be charted separately.

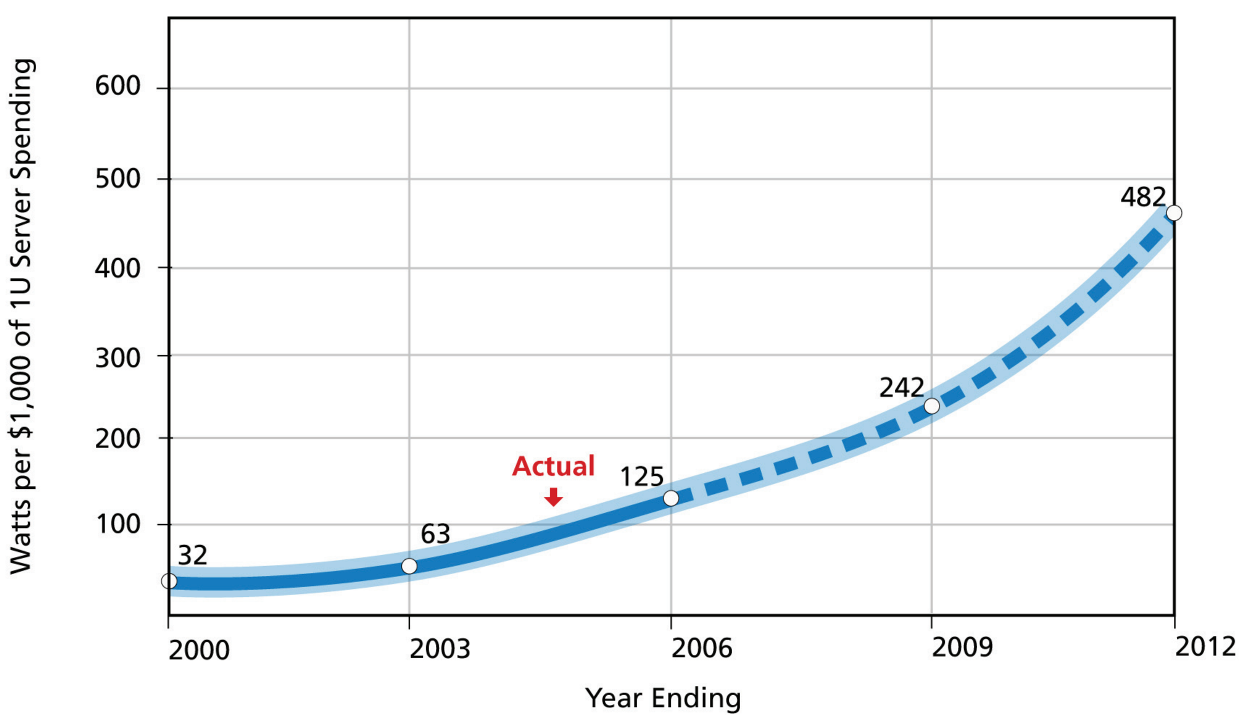
Spending Chart 1 shows Embedded Watts per $1,000 of 1U Server Spending (not including licenses, software, server maintenance, installation, and other IT costs). The *Institute* is refining this chart (for a subsequent revision of this white paper) to include performance units and invites additional data. The important point of this graph is not the absolute value of the numbers, but that the trend line is continuously moving up. Until this upward line is flattened or reversed, site TCO will continue to escalate, invisibly consuming IT funds which would otherwise be available for worthy projects that would contribute to enterprise productivity.

Author's Caveat: Healthy Skepticism Warranted

The energy-efficiency arguments presented thus far (i.e. server performance 3X, energy efficiency 2X, and a price decline of 0.94x every two years) would predict a much more radical rate of growth in data center power consumption than is actually happening. Koomey reports a doubling in consumption between 2000 and 2005, which is significantly lagging what might be expected. Careful variance analysis of Koomey's numbers indicates that 62 percent of the consumption increase was purely due to a greater number of servers being installed and only 38 percent was from increasing Watts per server. This conclusion is counter to the *Institute's* earlier arguments. *Institute* member data for 16 major sites with 29,000 kW of UPS load during the 2000 - 2005 time period shows consumption went up by 53 percent with no increase in floor space. (During this period, site capacity increased, but at a much slower rate than consumption. As a result, load as a percentage of site capacity went from 43 percent in 2000 to 58 percent in 2005. Many of these companies currently have active site infrastructure capacity expansion projects). How can we explain such significantly slower growth rates?

Data centers are a stew of non-homogeneous hardware technologies from many different manufacturers. Servers are an important component of total data center load, but so are mainframes, disk and tape storage, networking, and communications equipment. Consolidations of applications are likely to have played a significant role in reducing what otherwise would have been major

Spending Chart 1—Embedded Watts per $1,000 of 1U Server Spending



additional growth in power consumption. Something is happening at the macro level of the data center which is moderating power consumption growth below what has been argued earlier in this paper. Until this factor (or factors) is better understood, readers should exercise some skepticism about future power growth predictions. However, users can't go back to sticking their heads in the sand and hoping the trends being reported will go away. Doing so would be unrealistic and irresponsible. Whether data-center power consumption over the last five years is by up 53 percent or by 100 percent, the magnitude of these numbers is just too large to ignore. The remainder of this paper is devoted to things managers and senior executives should be doing to prudently manage this consumption growth to assure a positive return on investment.

C-Suite Beware: Sub-Optimal IT Investment Decisions will Result if the Cost-per-Compute-Unit Fails to Include Site TCO

A classic story within the *Institute's* Site Uptime Network is that of a company that invested $22M in blade servers without including the site facility representatives in the decision.

Only after the blades were ordered were the facilities executives consulted. Just to install the blades, a site infrastructure upgrade investment (site CapEx) of an additional $54M was required. For three years of operation, another $21M was required for site OpEx.

The true three-year cost of the decision to install blades was $97M, the majority of which was site CapEx and OpEx not considered in the original investment analysis. The point to this story is not that a bad decision was made, but that significant costs in the ROI calculation were overlooked because data center facilities professionals were not included in the blade investment decision.

Falling cost per computational unit is the result of Moore's Law and has fueled productivity growth in the US and global economy for the last 40 years. At this point, however, other whole-system constraints such as power consumption and weight may begin to place limits on future IT growth. The total weight of a full rack of servers can now exceed that of a large automobile, only in a much smaller footprint, and many buildings do not have sufficient structural strength to support this much weight if large numbers of cabinets are to be densely installed. This is not the first time Moore's Law has come up against physical constraints and an imminent breakdown was forecast. One of the purposes of this paper is to call industry attention to the whole-system constraints, including user economics, that computing is now encountering in the hope that solutions will be quickly found to the problems identified herein.

Other IT costs are falling (such as the ratio of system administrators per 100 servers and operating system licensing costs). These efficiencies may more than offset growing site TCO, but, to avoid being blindsided, senior executives should be including site TCO in their cost per compute unit calculations. *Table 2* shows the cost per compute unit dropping from $1M in 2000 to $1K in 2012. This is the historical 1,000-to-one ratio that has been going on for many years, but this is true only if hardware alone is included in the calculation. Before 2000, site TCO was not a significant factor and could safely be ignored. This is no longer true. If hardware and three-year site TCO are correctly included, the cost per compute unit in 2012 is not $1K, but $11K or 11X greater. A cost factor change of this magnitude typically makes a very big difference in any investment decision.

Table 2 – Three-Year Site TCO per 1U Compute Unit

		Cost Per Compute Unit	
Year	1U Compute Units Per $1.0M (Year 2000 = 1)	Hardware Only	Hardware + 3-year Site TCO
2000	1	$1K	$1,460K
2003	5	200	380
2006	27	37	104
2009	140	8	37
2012	729	1	11
Improvement	729X	1000X	133X

Because server compute performance is increasing while hardware prices are falling, a single computational unit of 1U performance that cost $1M in year 2000 would cost just $1K in 2012. This 1,000X reduction is dramatic, but is only true if just server acquisition cost is considered. When enterprise server and three-year site TCO costs are included, the cost reduction per compute unit is much less pronounced, dropping by a factor of 133, not 1,000.

Data Center Power Consumption and Carbon Footprint

In 2000, data centers (servers, storage, tape, networking, and site infrastructure) consumed approximately 0.8 percent of total US electrical consumption. In 2005, despite a seven percent growth in electricity production, data center power consumption grew to consume approximately 1.4 percent of the total. Even assuming some utilization gains from virtualization and incremental improvements to server efficiency, data centers are still projected to consume 2.3 percent of total electricity production by 2010[5]. Some estimates place it even higher.

After 2010, when the one-time benefits of virtualization are likely to have been fully assimilated, the trend of doubling power use every five years will potentially re-assert itself, and the absolute numbers will represent needed construction of multiple thousand-megawatt coal or nuclear power generating plants.

[5] Based on total US retail electric energy (kWh) sales, from EIA's Electric Power Annual 2006 and Annual Energy Outlook 2007. Data center consumption numbers are based on the EPA's report to congress on server and data center efficiency. The Institute believes that energy used by storage could be substantially higher than that cited by the EPA, meaning that data centers could, in fact, account for a slightly larger fraction of US electricity use than we report here.

Accompanying this growth are the obvious concerns about pollution, carbon emissions, and environmental sustainability.

This ongoing growth in electrical consumption is already attracting governmental policy and regulatory attention to IT. Even more attention can be expected in the years just ahead. Companies setting goals for reducing their carbon footprint cannot afford to ignore data centers, which (with the exception of heavy industries like steel and aluminum) are the most common and concentrated consumers of energy in the US economy.

It is fortunate that IT can argue that production of compute cycles in the data center in turn drives user business applications that may save vast amounts of energy in other parts of the enterprise. As a result, even though data centers are major power consumers, at the enterprise level, they save energy. IT executives will have to perfect these arguments with true cost justifications and ready examples. In addition, boards of directors' mandates for carbon footprint reductions and the rapidly growing site costs outlined in *Table 1* show that reducing energy consumption is not just about being environmentally responsible, but it is also smart business. Efficiency for profit will provide strong motivation for the entire IT industry to quickly become much more energy efficient while also being good corporate citizens at the same time.

Long- and Short-Term Initiatives for Significantly Improving Data Center Energy Efficiency and Productivity

The long-term solution for restoring the economic productivity of IT is to have the rate of energy efficiency increase equal or exceed the rate of computational performance increase.

Increasing the rate of improvement in energy efficiency will require changing the technology roadmaps for chips, drives, and other IT hardware components. Clearly, the importance of energy efficiency has been elevated in R&D departments, but reducing power consumption growth to a net zero or less is far beyond what many technologists are currently thinking.

C-suite user executives should be demanding this fundamental shift in R&D thinking from their vendors. However, it will take at least five years, and perhaps a decade for this innovation to result in energy efficient products users can buy for their data centers. (Also, even if equipment could be made more efficient tomorrow, it would still take a few years for the existing installed stock to turn over.) In the meantime, the *Institute* has identified concrete actions users can implement to reduce energy consumption at little cost—just by eliminating current waste.

For an overview of these energy-saving techniques, please see the white paper that resulted from the 2006 Symposium titled *High-Density Computing: The Path-Forward 2006*. This document contains the *Five Gold Nuggets* that can reduce data center energy consumption by up to 50 percent. In summary, these are:

- Server virtualization, consolidation, configuration
- Enabling server power-save features
- Turning off servers no longer in use
- Pruning bloated software
- Improving the Site Infrastructure Energy Efficiency Ratio (SI-EER, *which the Institute is currently working to re-cast in more intuitive and technically accurate terms.*)

Likewise, the *Institute* is evolving its Integrated Critical Computing Environment™ (ICE) methodology and collaborative team management best practices for implementing effective energy efficiency measures while ensuring "24-by-forever" reliable uptime availability. To assist ICE Teams and other IT and site infrastructure technologists, the *Institute* is developing procedures for chartering effective ICE Teams and a set of metrics and methodologies for measuring, benchmarking, and improving data center energy efficiency.

The *Institute* is currently identifying the three to five unique performance factors cutting across traditional IT and facilities boundaries to define and determine a green data center. Once correctly defined, these factors will allow current energy optimization problems to be broken down into separate chunks to be solved by different groups and constituencies. Some of these people will be within user organizations, others will be in chip, storage, and hardware manufacturers, and yet others are in the site infrastructure design/engineering consulting community. Optimization of each factor will result in optimization of the whole system[6].

References

Aperture Research Institute. 2006. "Organizations Struggle with Data Center Capacity Management." www.aperture.com/about/aperture_research_institute.php

Belady, Christian. 2007. "In the Data Center, Power and Cooling Costs More Than the IT Equipment it Supports." Electronics Cooling. Vol. 23, No. 1, February 2007.

EIA. 2006. Electric Power Annual, 2006. www.eia.doe.gov/cneaf/electricity/epa/epat7p2.html

EIA. 2007. Annual Energy Outlook 2007. www.eia.doe.gov/oiaf/aeo/pdf/0383(2007).pdf

Koomey, Jonathan. 15 Feb 2007. "Estimating Total Power Consumption by Servers in the U.S. and the World." Analytics Press. enterprise.amd.com/Downloads/svrpwrusecompletefinal.pdf

U.S. Environmental Protection Agency, ENERGY STAR Program. 2007. "Report to Congress on Server and Data Center Energy Efficiency (pursuant to Public Law 109-431)."

[6] *The thesis covered in this review-and-comment draft forms the principal philosophical basis for the upcoming Institute design Charrette titled:* Data Center Energy Efficiency by Design: Engineering & Managing Site Infrastructure Systems of the Future*, October 28-30, 2007, Santa Fe, NM. For information on participating in this forward-looking data center design project, visit: http://www.uptimeinstitute.org/charrette.*

White Paper Reviewers

Each subsequent draft of this paper is being publicly reviewed and will continue to evolve as an open-source document. We appreciate all feedback. However, many constructive comments came in too late to be incorporated in this release, but all reviewers will receive a response. The next draft in early August, 2007 will include an addendum of the most thoughtful and helpful critiques that further the discussion of this complex issue. One important criticism that could not be addressed in this version was that the information was too technical for C-suite executives. The *Institute* will solve this problem by creating a separate *executive summary* specifically for C-suite consumption. Among the reviewers providing comments on this white paper were Stan Burchell, Abbott Laboratories; Lane Pinnow, Agilent; John Menoche, APC; Michele Blazek, AT&T; Allen Beatty, AT&T; Leslie Ingram, Barclays; Gregg Balzer, Blue Cross Blue Shield of Florida, Inc.; Andy MacMillan, Bureau of Land Management (US Dept. of the Interior); John Strothman; Strothman Associates, Inc, ; David Moss, Dell; John Pflueger, Dell; William Gilliland, Digital Business Consulting, LLC; Kenneth Uhlman, Eaton; Rick Sawyer, EYP Mission Critical Facilities; John Sears, Hitec Power Solutions; Christian Belady, HP; Dave Rotheroe, HP; Ken Baker, HP; Larry Rushing, HP; Robert Hughes, HP; Wade Vinson, HP; Henry Wong, Intel; Herb Villa, Rittal; Amory Lovins, Rocky Mountain Institute; Steef van Schaik, Shell; Rich Frey, SIAC; Ron Kalley, State Farm Insurance; John Diamond, Strategic Facilities, Inc.; John Dennis, Sutter Health; Ernest Meyer, Texas Instruments; Charles Hobart, Washington Mutual; Dave Ellis, Vanderbilt University; Dave Cosner, University of Maryland; Lance Windham, USAA; Edmund Orchowski, Verizon; Mike Gagnon, Wright Line.

About the Author

Kenneth G. Brill is the founder and Executive Director of the *Institute* and the 100-corporate member Site Uptime Network in both North America and Europe. He holds an undergraduate degree in electrical engineering and an MBA from the Harvard Business School. Many industry innovations including dual power and the industry's Tier system for evaluating data center design level trace back to his original conceptual work. In 1999, recognizing that increasing heat density would become critical to IT availability; Mr. Brill worked closely with the Thermal Management Consortium to publish the first 2000-2010 Heat Density Product Trends white paper. This foundational document predicting many of the problems now facing IT executives was updated in April 2006. He has recently contributed to original research that has now shown data centers consuming 1 percent of total US electrical power production in 2000, growing to 2 percent by 2005, and predicted (based on different assumptions) to grow to 3-4 percent by 2010 (these figures are higher than found in other documents because Brill contends others have significantly underestimated disk storage power consumption). The absolute numbers caused by this growth has raised serious questions about the adequacy of the US electric supply in an era of increasing concern about carbon emissions. These issues have caught the attention of Congress. Public Law 109-431 signed December 21, 2006 requires that the Environmental Protection Agency (EPA) report to Congress by June of 2007 (now extended to late July) on these issues. Mr. Brill has been providing technical support and leadership to this process. He has authored or contributed to many white papers and is a frequent commentator in business and technology media on data center and site infrastructure design, engineering, and management issues. His current focus is on strategic and business impacts of The Economic Meltdown of Moore's Law.

About the Uptime Institute, Inc.®

Since 1993, the Uptime Institute (*Institute*) has been a respected provider of educational and consulting services for Facilities and Information Technology organizations interested in maximizing data center uptime. The *Institute* has pioneered numerous industry innovations, such as the Tier Classification System for data center availability, which serve as industry standards today. At the center of the *Institute's* offering, the 100 global members of the Site Uptime Network® represent mostly Fortune 100 companies for whom infrastructure availability is a serious concern. They collectively and interactively learn from one another as well as from Institute-facilitated conferences, site tours, benchmarking, best practices, and abnormal incident collection and analysis. For the industry as a whole, the *Institute* publishes white papers, offers a Site Uptime Seminar Series and a Site Uptime Symposium and Data Center Design Charrette Series on critical uptime-related topics. The *Institute* also conducts sponsored research and product certifications for industry manufacturers. For users, the *Institute* certifies data center Tier level and site resiliency.

Uptime Institute
2904 Rodeo Park Drive East, Building 100, Santa Fe, NM 87505
505.986.3900 · Fax 505.982.8484 · uptimeinstitute.org

 TUI3008A