Minimizing Energy-Related Costs **In Data Centers and Network Operations Centers**

Energy Management "By Design"

by Barry J. Needle

AGO Research

energy management by design: *to minimize energy consumption and maximize energy reliability in data centers and network operations centers by leveraging a thorough understanding of the "as designed" functioning of the IT and operations support infrastructure.*

Minimizing energy-related costs in the data center and network operations centers is a hot subject in the it world today, and for a number of good reasons, not the least of which are that energy is an increasingly costly resource and idle computing equipment costs businesses dearly. Usage continues to rise at a seemingly inexorable rate. The amount of energy being used is staggering when you consider it is being used ultimately to switch and transport the bits of information…the ones and zeros of our lives for which we have an apparent insatiable desire, and regard much as a utility like water, gas, and electricity. And, when the bits stop, the cost ranges from pennies of annoyance to millions of real dollars of lost revenue or business disruption.

The challenge to computing facilities managers and owners is to reduce operating costs, by minimizing energy consumption and costly downtime, while maximizing power reliability and IT application performance and availability. Lately, in an effort to understand and improve the energy efficiency of IT enterprises, the industry has adopted the metrics, Power Usage Effectiveness and Data Center Infrastructure Efficiency, proposed by The Green Grid organization[1]. These metrics, via Facility and IT Equipment Power ratios, facilitate more informed decisions on measures to be taken to reduce the total cost of operating data centers while managing increased service demands. Functions such as data center commissioning, and testing schemes on system modifications designed to address how much risk is prudent to accept in the quest to save energy, benefit from the information revealed by the power ratios. Further, these PUE and DCiE allow individual operators to measure the effectiveness of efficiency improvement programs by comparing the energy efficiency of their facilities to those of other like organizations

At the core of all actions to improve energy efficiency, reliability, and availability is the answer to the question of cause and effect. Any change to the operation of a data center will have an effect on the amount of power the IT equipment uses, the quality and stability of the power distribution system, and the health of the computing facilities output. Moreover, since every data center is unique design-wise and operationally, then the actions taken to manage energy at any data

---

[1] The Green Grid, "The Green Grid Data Center Power Efficiency Metrics: PUE and DCiE" December 2007

center will be most effective if they are formulated specifically for that site. The Paladin® Live power analytics platform from EDSA Micro is the only software capable of producing the focused information to guide a site-specific energy management program – energy management by design.

The Data Center Powering Crisis

The digital electronics systems which shuttle our 'infobits' are located typically in buildings, which because of the high power requirements of the IT equipment, and the supporting power and cooling infrastructure, are currently up to 40 times more energy intensive than a typical office building. The data center is more like an industrial complex with respect to energy usage.

According to a recent EPA report[2], the power demand of the data centers in the U.S. is significant…and growing.

» The energy consumption of servers (including cooling and auxiliary infrastructure) in U.S. data centers has doubled in the past five years and is expected to almost double again in the next five years *[2011]* to more than 100 billion kilowatt-hours (kWh), costing more than $7.4 billion annually (2005 dollars). "The peak load on the power grid from these servers and data centers is currently estimated to be approximately 7 gigawatts (GW), equivalent to the output of about 15 baseload power plants. If current trends continue, this demand would rise to 12 GW by 2011, which would require an additional 10 power plants."

» Data centers consumed about 60 billion kWh in 2006, roughly 1.5 percent of total U.S. Electricity consumption.

It s no mystery to the data center operator that most of the energy to the servers ends up as heat. The cost to power and cool racks of installed servers is significant, and is forecasted to become greater relative to new server spending. The Figure 1 below graphically shows the increasing proportion of power and cooling relative to server spending worldwide. The data shows that for each dollar spent on a new server in 2005, forty-eight cents was spent on power and cooling. This is more than twice that ratio in 2000. In 2010, it is projected that this ratio will rise to $1:$0.71

---

[2] EPA, "Report to Congress on Server and Data Center Energy Efficiency", August 2, 2007

And it isn't just the expense to power and cool servers. In fact, there is potentially an extreme price to pay for no power … idle servers. Unintended downtime is costly. What is the average price? A million dollars an hour. That's what IT system downtime costs American business, according to a keynote address by the META group (now Gartner, Inc.) given six years ago. And the cost hasn't come down.

Worldwide Expense to Power and Cool Installed Server Base, 1996-2010
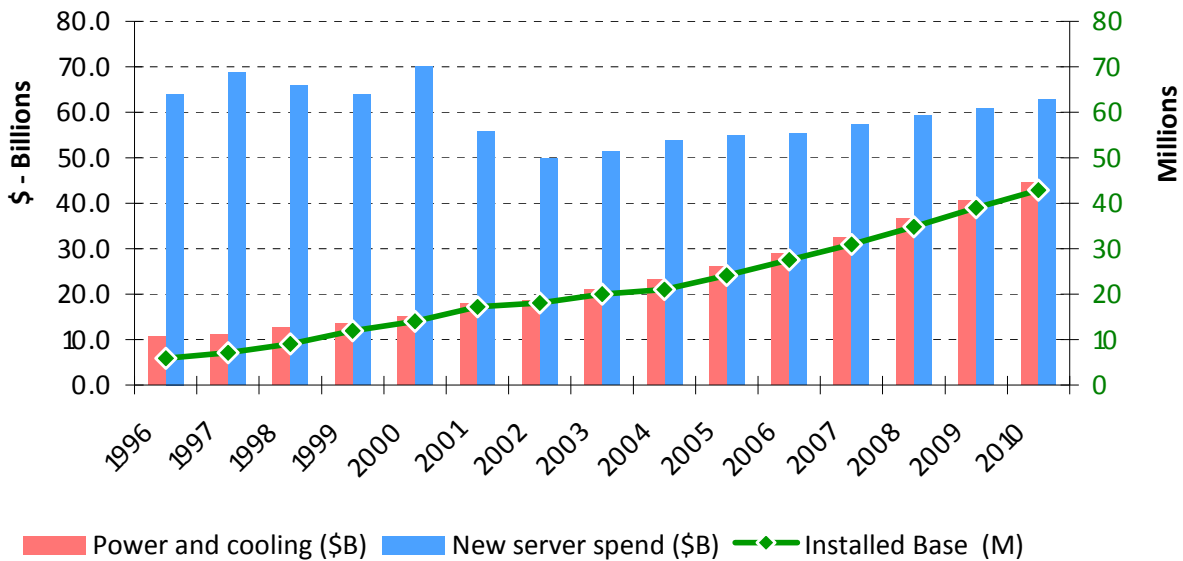

Power and cooling ($B)   New server spend ($B)   Installed Base (M)

Figure 1

Source: IDC, 2007

It should be noted that the forecasted power demand cited previously does not reflect unmitigated historical growth extrapolations. During the last several years, the industry's attention to the growing crisis of unbridled energy demand has fostered many of the ideas generating the positive effects of current energy-efficiency trends.   According to the EPA report, however, there remains significant potential for further improvement in reducing future energy demands, and realizing notable, environmental "green" effects. Three energy-efficiency scenarios were developed to explore the impact of technological approaches that could be deployed without unacceptable risk to data center performance, reliability, and availability.

Comparison of Project Electricity Use, All Scenarios 2006-2011

← *Historical energy use* | *Future energy use projections* →
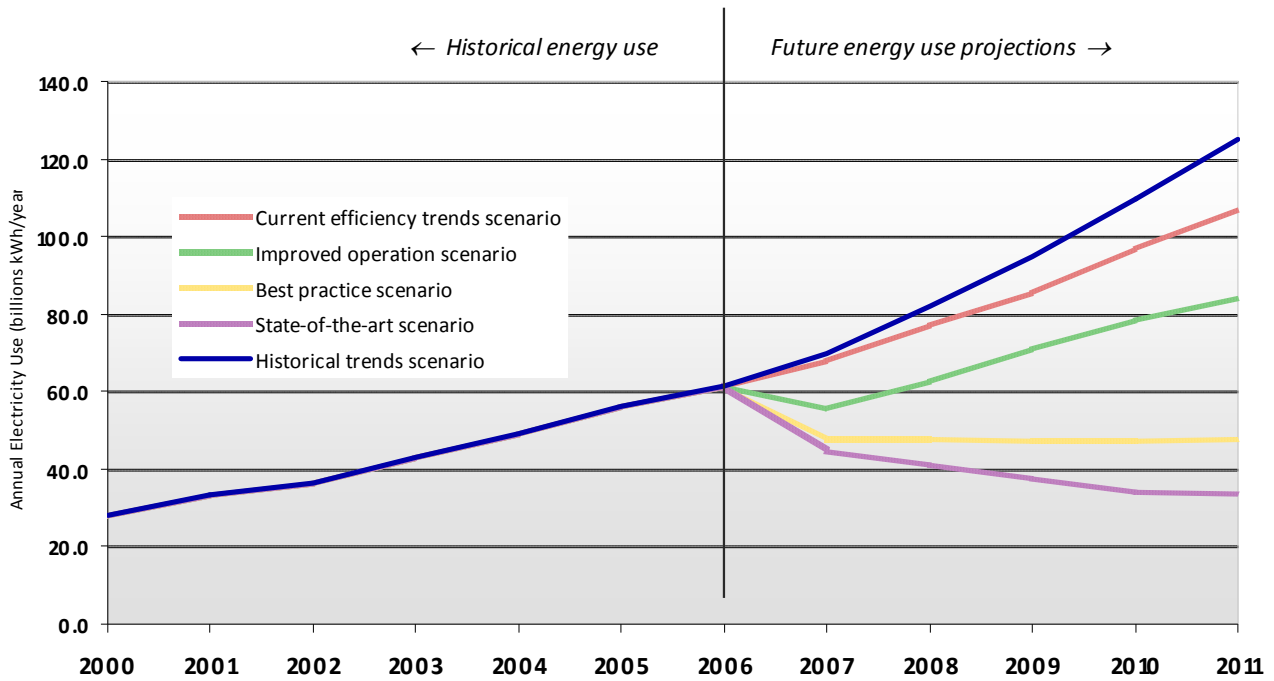
Dubbed "improved operation", "best practice" , and "state-of-the art," the envisioned improvements in energy efficiency would be significant resulting in a potential dramatic reversal of energy demand trends. The accompanying reduction in the total carbon footprint (green house gas emissions) from the operation of IT facilities is noteworthy given the greater than 21% contribution[3] to total greenhouse gas emissions from power plants. The annual savings in 2011 ranging from approximately 23 to 74 billion kWh is compared to the current efficiency trends scenario. Annual electricity costs would be reduced by $1.6 billion to $5.1 billion.

Projected CO_2 Emissions Associated with the Electricity Use of U. S. Servers and Data Centers (MMT -CO_2/Year), All Scenarios, 2007 to 2011

Scenario	2007	2008	2009	2010	2011	2007-2011 Total	% of current efficiency trends scenario
Historical Trends	44.4	51.2	59.2	69.2	78.7	302.8	111%
Current Efficiency Trends	42.8	47.9	53.6	60.5	67.9	272.8	100%
Improved Operation	34.8	39	43.5	48.4	53.1	219	80%
Best Practice	30.2	30	29.8	29.7	30.1	149.8	55%
State-of-the-Art	28.1	25.7	23.5	21.4	21.2	119.9	44%
							- EPA, 2007

---

[3] Emission Database for Global Atmospheric Research version 3.2, Fast Track 2000 Project. This value is intended to provide a snapshot of global annual greenhouse gas emissions in the year 2000

The projected savings in electricity use correspond to reductions in nationwide carbon dioxide (CO_2) emissions of 15 to 47 million metric tons (MMT) in 2011.

Breakthrough in Data Center Design?

A notion, not often considered, is that a data center's consumption of energy is <u>by design</u>. The design and specification process, from intended server loading, to IT equipment selection, and the power and cooling infrastructure, determines the center's power demand. All the engineering done for detailing the support infrastructure (including power distribution and backup, server and electrical room cooling, and lighting systems), is to ensure the reliable operation and maximum uptime of the installed IT engine.

The data center's initial design is a snap-shot…once built, modifications are tested by trial and error. A company's computing services are far from being a static environment. Data center managers are called upon to implement new application loading schemes resulting in server consolidations and virtualization. Growth means new, energy-efficient equipment replacing old, thin provisioning, and capacity expansion assessments. The "error" part of testing is risky and potentially costly. And today, all changes are being audited for energy efficiency, risk of system instability and cost impact. One way or another, minimizing the power demands (cost of operation) of data centers while maximizing reliability and availability is the goal.

The issue is how to ensure system reliability and uptime, while managing power usage and risk of system upset cost-effectively. At a recent industry conference, the question of data center energy efficiency by design was considered with the hope that the IT industry might discover a technological breakthrough which would radically alter the design of data centers and economics of energy use. To the chagrin of the participants, no dazzling, new answers were forthcoming. Rather, the conclusion was that data center designers and owners needed to "tune up what they own," since no sea change solutions were in the offing.

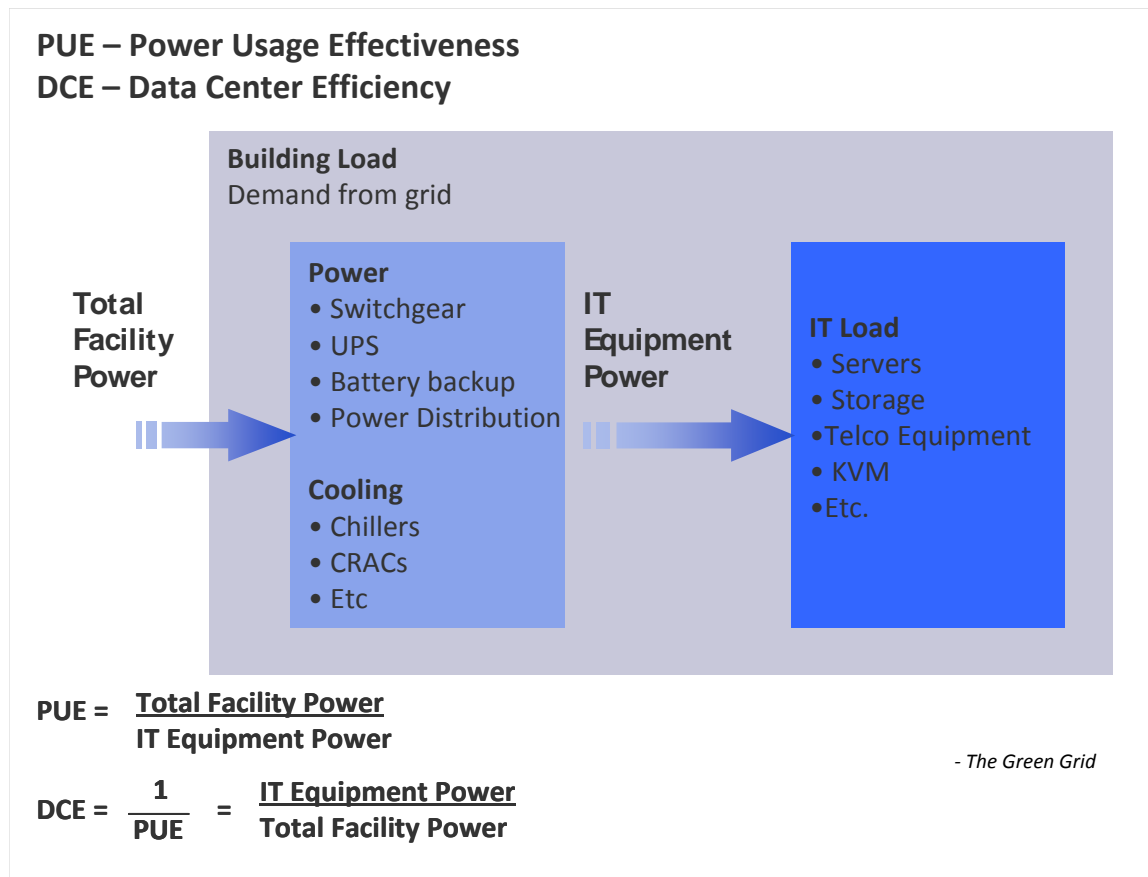**The Energy Efficiency Tune-up Tool - Energy Management By Design**

The Uptime Institute Design Charrette 2007[4] found a number of conditions which were necessary for the tune up.

> » A way to generate the metrics that show what performance levels can be reached and at what level are systems performing currently. "This granular benchmarking drives the tune-up process."

> » Thorough knowledge and operational experience of a specific data center.

---

[4] Executive Director Report: The Findings of the 2007 Charrette, Kenneth G. Brill, DESIGN CHARRETTE 2007, Data Center Energy Efficiency By Design

> » Practical engineering and economic analysis training with implementation skills focused on reducing risk of unintended downtime or reliability issues.

These conditions are met or facilitated by the energy management by design approach based on the Paladin® Live Real-Time Power Analytics software platform. As mentioned previously, the application of the PUE and DCiE power ratios are giving operators much better understanding of their facilities power and cooling performance from an energy demand perspective. By partitioning the IT Equipment Loading (servers, storage, network management, communication, etc.) from the Total Facility Power (switchgear, UPS, power backup systems, chillers, CRAC, etc.) demand, the data center manager gains a much clearer picture of the facilities performance dynamics. Paladin® Live is uniquely capable of deriving real-time, system-specific performance data.

PUE – Power Usage Effectiveness
DCE – Data Center Efficiency

Building Load
Demand from grid

Total Facility Power

Power
• Switchgear
• UPS
• Battery backup
• Power Distribution

Cooling
• Chillers
• CRACs
• Etc

IT Equipment Power

IT Load
• Servers
• Storage
• Telco Equipment
• KVM
• Etc.

- The Green Grid

$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

$$DCE = \frac{1}{PUE} = \frac{\text{IT Equipment Power}}{\text{Total Facility Power}}$$

Site-specific data center design information, powering the Paladin® Live engine, provides the basis for exclusive insight into operations. The powerful, concurrent simulation capabilities of Paladin® Live via its "Blackboard" feature allow virtual testing of planned modifications, and observation of their impact on the PUE and DCiE power efficiency metrics without the risk of potential downtime.

A closer look at Paladin® Live

When installed, the power and uniqueness of the platform is derived from the complete encoding of the design specifications from the original, as-built power infrastructure. All power system electrical parameters are calculated from the stored design specifications, and, during the data center's normal operation, compared with the real-time power data. At any time, Paladin® Live can accurately corroborate as-specified power parameters, determine if there are system anomalies, and predict when and where there are potential vulnerabilities for system and equipment failure. Further, the Paladin® Blackboard™ feature allows users to capture current system state data, and run detailed "what if" simulations to verify system operations for the data center commissioning process, to investigate the effects that equipment rearrangement, configuration modifications, capacity expansion and other data room modifications might have on the live system without the risk of actually doing live testing. Simulations of maintenance and repair actions can help discover unforeseen program vulnerabilities and guide optimum cost-effective scheduling. Facilities engineers can review powering schemes for reliability and capacity. IT managers, concerned with availability and service level agreements, can explore dynamic application loading scenarios in a virtual environment without the risk of unintended downtime.

What Is Site-Specific Information Worth?

Beginning with the data center commissioning process, the "as designed" insight into a data center's electrical power infrastructure, and the ability to simulate in the entire power distribution system in a virtual environment, will reduce the overall process costs by generating and maintaining a knowledge of the physical infrastructure, verification of performance, and ability to probe for potential out-of-specification system parameters.

Commissioning costs are high, but the return on the investment is significant. According to a report by Einhorn Yaffee Prescott (EYP), a global consulting engineering firm, the data center owner should plan to spend 1 to 2% of the overall data center project cost on commissioning. In most cases, the owners will see a 5-10% ROI benefit in terms of overall data center performance as a result of commissioning[5]. With today's cost of data center construction approaching $2,500 per sq. ft., the commissioning of a 50,000 sq. ft., Tier IV facility will run close to $2,500,000. Paladin® Live has the potential to save 10 to 25% of the overall commissioning process costs.

Downtime will impact business revenues. While intended downtime can have a minimally disruptive, low cost impact on the normal operation of a business,

---

[5] Einhorn Yaffee Prescott, Data Center World, *Everything You Need to Know About Commissioning*, March 2006

unintended downtime has the opposite effect. Depending on industry sector the value of one hour of unavailability can range from about $100,000 to more than $6,000,000. The real-time power analytics capability of Paladin® Live can intelligently predict the timing and location of potential system upsets, and, in the case of a downtime episode, can quickly apprise the right people as to the cause and solution. Since time is money, reducing overall downtime by as little as six minutes per year can mean a potential savings of about $100,000 if downtime is worth $1 million per hour. Paladin® Live can formulate truly predictive diagnostics based on system design boundaries, and the implications of variable operating conditions from system aging. Intelligently scheduled system maintenance or repair based on a reliability assessment rather than a simple periodic basis can be less upsetting and costly. The Blackboard feature can be used to explore the impact of various maintenance or repair schemes without live testing.

Key to an effective energy management program is accurate information regarding the consumption of energy. Based on the amount of IT equipment in racks, the power distribution and cooling equipment infrastructure, and the variations in application loading, Paladin® Live can report accurate, real-time energy usage. This data can be compared to the "as-designed" energy usage calculated by Paladin® Live to give insight into system unbalances, capacity restraints, or overloads. The results of virtualization and other energy efficiency measures can be followed and assimilated. Paladin® Live can suggest scenarios for improved energy utilization based on its predictive diagnostics ability and by "what-if" simulation. At the current energy costs (~$0.089 kWh), a nominal realized annual savings of ten percent for even a relatively small, lightly-loaded data center is significant – greater than $100,000.

As mentioned before, computing facilities engineers are generally focused on the reliability and capacity of the data center's power distribution system; the data center manager is concerned with server availability and service level agreements. While they may be preoccupied with different aspects of data center operation, they both are in agreement regarding taking risks. They do not want to take them. The adage, "no pain, no gain" is simply not part of their conversation. The fact is, though, energy conservation schemes, thermal efficiency advances, capacity improvements, server loading rearrangements, new technology applications, and other energy management measures involve the risk of unintended consequences. The simulation of a system's performance in a virtual environment is the safest way to test a system modification and assess risk. Paladin® Live is a uniquely powerful platform on which to test system changes. The platform maintains real-time awareness of the actual, in-service operations, while the Paladin® Blackboard™ holds a mathematical clone of the operations in a virtual environment. In a manner analogous to recording a film, the Blackboard records the data of the live environment frame-by-frame, enabling a user to analyze a chronological image of the live operation. "What if"

simulations of the effects of any changes on the actual configuration of operating parameters can be done "virtually" before testing them on a live system, and the high costs of unintended system failure or performance degradation may be avoided.

Return on Investment – Paladin® Live

Is site-specific information worth the investment in the Paladin® Live platform? On the basis of the current costs of data center commissioning, the average cost of electric power, realized savings of power, and minimized downtime, the investment is readily justified. An example of an ROI calculation for a Tier III data center is shown below.

DATA CENTER ENERGY CONSUMPTION

Data Center Size, sq. ft.	50,000	
New Data Center Commissioning Costs, assume 10% of 1.5% of $2000/sq.ft. construction costs	$0.15	
Heat Loading, watts/sq.ft.	40	100
Monthly Energy Costs @ 8.9 cents/ kW-hr	$0.28	$0.70
Realized Annual Energy Savings, assume 10.0%	$0.34	$0.85
Total Annual Downtime Reduced by 6 minutes at $1,000,000 /hour	$0.10	

PALADIN LIVE INSTALLED COSTS

New Data Center at $6.0/sq.ft.	$0.30	
Existing Data Center at $12.0/sq.ft.	$0.60	

ROI , YEARS

New Data Center	0.51	0.27
Existing Data Center	1.37	0.63

All $ in millions

	- less than 6 months
	- 6 to 12 months
	- more than 12 months

Notes:

1. The electricity cost is based on a constant rate of $0.089/kWh including both energy and kW-based demand charges. This rate is reflective of Site Uptime Network members information and the national average. Many regions pay significantly more ($0.14 for New York City).

2. 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. [Uptime Institute].

3. The cost of commissioning is based on information from *"Data Center Projects: Commissioning",* a white paper by Paul Marcoux of APC. The cost benefit of using Paladin® Live during the commissioning is estimated to range to 25%. 10% is used for this illustration.

4. Nominal heating loading data, 40 and 100 kW/sq. ft., are used for this example. The greater the loading, the shorter the ROI.

5. Expected savings on annualized energy costs is reported to range between 20-50% in some cases. 10% of the annual energy consumption for the 50,000 sq. ft. at the loadings shown is conservative.

6. Downtime costs are variable. The example here assumes the industry average of one million dollars per hour. A conservative, six-minute reduction in downtime is used for the illustration

7. The cost of installation for the Paladin® Live platform averages $6/sq.ft. for a new data center; about double for retrofitting an existing data center.

The payback for the Paladin® Live platform in this example is, at most ½ year for a new data center, about 15 months for the 40 W/ sq. ft. existing facility, and only eight months for the 100 W/sq. ft. loading. Time-based discounted cash flow analysis would allow greater insight into the investment.

Bottom Line

The concern for system availability was the exclusive guiding design concern for data centers and network operations in any mission critical element from the very beginning. The combination of Paladin® Live's ability to manage critical power from the forensic perspective, and from a proactive predictive perspective is unique. This is especially relevant because the Power Analytics approach which provides management metrics to simplify and demystify energy management are true enabling technologies.

The added complexity of energy management will increasingly drive system and financial decisions. The Paladin® Live product addresses the continuum of energy management from availability and performance to reliability and quality; a timely and powerful solution for the 21st Century technological enterprise.