14 AUG 2011

Add new comment

Managing Power Usage with Energy Efficiency Metrics: The Available Metrics and How to Use Them

Rate this item (1 Vote) font size

Data centers consume an enormous amount of power. According to a detailed study conducted by the U.S. Environmental Protection Agency (EPA) in August 2007, data centers in the U.S. consumed 61 billion kilowatt-hours, or 1.5 percent of the nation's total electricity, in 2006, adding some 40 million tons of CO2 to the atmosphere. That amount is twice what was consumed just six years earlier, and the EPA forecast data center power consumption to double again from 2006 to 2012—a milestone that may have been reached already, according to some estimates.

Companies are now paying very close attention to power consumption throughout their data centers. To support this growing need, the industry has created both the metrics and tools required to measure and manage power utilization effectively. This article describes the various rating systems now available and discusses how each can be used to improve energy efficiency in any data center.

Before exploring the efficiency metrics, it is important to put power consumption in data centers into the proper perspective. The cost of electricity currently represents about 25–40% of the operational expenditures for the typical data center. Both the U.S. Department of Energy and Gartner have observed that the cost to power a typical server over its useful life can now exceed the original capital expenditure. Gartner also notes that it can now cost over $50,000 per year to power a single rack of servers.

This growing electrical demand might be acceptable if all of that power were being put to useful work. But that is not the case, as a considerable amount of power (even more than half) continues to be wasted, with the major cause being underutilized servers. And the growing demand for server capacity (increasing by an estimated 11% annually) ultimately threatens to exceed the total power and/or cooling capacity available in many data centers, forcing costly upgrades or expansion. It is this threat of outgrowing a data center that has motivated organizations to begin monitoring and managing power utilization.

Getting Started with Energy Efficiency Metrics

One common means for assessing data center energy efficiency is the *Power Usage Effectiveness (PUE)* rating system. PUE, created by The Green Grid, is the ratio of total power or energy consumed and the power or energy used by the IT equipment. Version 2 of the rating system (PUE 2) permits measuring either power (in kilowatts) or energy (in kilowatt-hours), with a strong preference for the latter. The typical data center today achieves a rating of about 2.0; in other words, only half of the total energy consumed is being used by the IT equipment (servers, storage and networking infrastructure), with the other half going to the

cooling, lights and the inherent inefficiencies in power distribution systems. Another common rating system is *Data Center Infrastructure Efficiency (DCiE),* which is the reciprocal of PUE: the ratio of the power used by the IT equipment to the total power consumed.

To ensure meaningful PUE or DCiE measurements, it is important to collect data often. Frequent measurements are necessary to capture changes in power utilization at different times of year, different times of the day, and most importantly, during periods of peak and low demand. It is also useful to collect information on environmental conditions at the same time, such as humidity, and hot and cold aisle temperatures.

A Data Center Infrastructure Management (DCIM) system is a powerful tool for collecting actual and frequent measurements within a data center. The DCIM system handles all of the measuring, calculating, averaging and comparing automatically, enabling the organization to establish a baseline and to determine the progress toward meeting energy efficiency goals.

The typical DCIM system supports both the industry standard and popular proprietary protocols used to measure power consumption, which means there are no special agents to install or extra wires to run to measure power at the building, circuit and device levels. Most DCIM systems also measure environmental conditions, such as temperature, humidity and airflow, throughout the data center. Some offer advanced capabilities like auto-discovery, capacity planning, building management system (BMS) integration, sophisticated yet intuitive dashboards, comprehensive reporting and more. The best ones provide real-time monitoring, advanced analytics and the ability to automate processes in cooperation with load-balancing or virtualization systems to continuously match server capacity with demand.

The more accurate the PUE measurements become, the more necessary it is to have a DCIM system. The reason exists in the granularity of the four PUE categories, and reporting PUE results now requires specifying which category is being used. The first and least-accurate category (PUE_0) measures IT power consumption (kW) at the output of the uninterruptible power supply (UPS), and it should be taken at the peak. The remaining three categories all measure the actual energy consumption (kWh) of the IT equipment at progressively more granular levels: PUE_1 at the UPS output, PUE_2 at the power distribution unit (PDU) output and PUE_3 at the IT equipment input. PUE reporting also just recently started to require actual data collection over 12 months to eliminate seasonal benefits.

Improving PUE/DCiE Ratings

The EPA has set a target for data centers in the U.S. of a PUE rating between 1.1 and 1.4 (or a DCiE of 0.9 to 0.7). The benefits of reaching this target range can be profound; for example, an improvement in PUE from 2.3 to 1.3 nearly doubles the power available for IT equipment. With today's typical data center having a PUE rating of around 2.0 (or a DCiE rating of 0.5), achieving this level of improvement will probably require a range of initiatives, including:

- Right-sizing the UPS and power distribution equipment to minimize inefficiencies
- Eliminating cooling inefficiencies, upgrading the computer room A/C system to allow for variable cooling and/or making greater use of outside air for cooling
- Adopting a hot/cold aisle configuration (which will involve optimizing the placement of servers in racks and rows) and increasing cold aisle server inlet temperatures to 80.6°F (27°C) as recommended by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE)
- Consolidating and virtualizing servers to improve overall utilization
- Refreshing IT equipment to newer, more power-efficient systems (see additional details below)

To help prioritize these measures, and perhaps some other potential power-efficiency initiatives, it is prudent to first establish a solid baseline for the existing data center. A DCIM system can quickly and accurately provide just such a baseline, and the results normally reveal those areas that will offer the most opportunity for improvement.

Improving server utilization through consolidation and virtualization can increase processing power while reducing energy consumption and cooling needs, as well as rack space utilization. But be aware of a possible surprise with seemingly counterproductive results in these situations: the improvement in IT asset efficiency can actually cause PUE to increase because less power is now going to the servers, while the overhead (cooling, lighting, power distribution, and so on) remains largely the same. But such good deeds should not be punished with these “false negative” ratings. This inherent limitation in PUE and DCiE has led to creation of some other rating systems that more accurately assess efforts to improve energy efficiency in data centers.

Improving IT Energy Efficiency

The reason PUE and DCiE can produce counter-productive results is that neither rating system takes into account IT energy efficiency. Indeed, using older IT equipment that consumes (and wastes) more power results in an “improvement” (lower PUE or higher DCiE).

To remedy this shortcoming, both McKinsey and Gartner have created more-comprehensive rating systems that take into account IT energy efficiency: *Corporate Average Datacenter Efficiency (CADE)* and *Power to Performance Effectiveness (PPE),* respectively. Both systems are designed to help IT managers address the primary source of waste in data centers today: underutilized servers. Average server utilization in the U.S. is 10% or less for dedicated servers and in the range of 20% to 30% for most virtualized environments.

CADE, for example, is the product of facility efficiency (in effect, DCiE) and IT asset efficiency, the latter being derived from a combination of the utilization and efficiency of the IT equipment. Organizations that have refreshed, consolidated and/or virtualized servers invariably achieve a significant improvement in the CADE rating, with PUE and DCiE getting either better or worse depending on the circumstances.

The best ways to improve CADE and PPE ratings are to utilize more energy-efficient servers and to increase overall server utilization. Organizations routinely refresh aging servers to take advantage of the improved performance made possible according to Moore’s Law. But determining the optimal time to refresh particular servers can be a challenge. Newer servers inevitably offer superior price/performance, but can their total cost of ownership (including power consumption as a major operating expense) be justified? And if so, which old servers should be replaced first, and by which new server model(s)?

Selecting the Most Power-Efficient Servers

To help IT managers make such choices, the EPA created an *Energy Star Rating System* for servers and other IT equipment. But Energy Star has a fundamental flaw similar to the one in the PUE and DCiE rating systems: it focuses on the power supply, ignoring the transactional efficiency of the server itself, and it does not factor in the year/age of the equipment. With an average doubling of server performance every two years, this is a serious shortcoming because there is no way to determine what an older server's Energy Star rating means relative to a newer server available today. As a result, to make a more fully-informed decision, IT managers must make estimates using conservative nameplate ratings and performance specifications, which can produce inaccurate and misleading results.

To address this shortcoming, Underwriters Laboratories (UL) created a new performance standard based on the *PAR^4 Efficiency Rating.* PAR^4 provides an accurate method for determining both absolute and normalized (over time) energy efficiency of both new and existing equipment on a transactions per second per watt basis. To calculate server performance using the UL2640 standard, a series of standardized tests is performed, including a Power On Spike test, a Boot Cycle test and a Benchmark test. The Power On Spike test characterizes the initial power surge that occurs when a device is first turned on or plugged in. The Boot Cycle test measures the time from the power off state until the device is fully functional and has stabilized at an idle state. The Benchmark test determines and classifies the server's overall performance and grants the PAR^4 certification for the year of production in one of the following categories: Green (most efficient), Gold (very efficient), Silver (fairly efficient) and Black (low efficiency).

These test results can then be used to determine idle and peak power consumption, along with transactions per second per watt and other useful annualized ratings. These metrics provide a very accurate means for IT managers to compare legacy servers with newer ones, and newer servers with one another.

Conclusion

IT and facility managers can no longer afford to ignore power consumption in the data center. Power is the limiting factor in the typical data center today, and extending the data center's life must now involve managing power more effectively. A capable DCIM system can establish the baseline and track progress toward energy efficiency goals using one or more rating systems (PUE, DCiE, CADE or PPE). DCIM can also help with other related tasks, including maximizing server utilization, reclaiming stranded power, optimizing equipment placement in aisles and racks, and what-if analyses for capacity planning, along with automated procedures to increase the reliability and energy efficiency of a data center.

About the Author

Clemens Pfeiffer is CTO of Power Assure and is a 25-year veteran of the software industry, where he has held leadership roles in process modeling and automation, software architecture and database design, and data center management and optimization technologies.

Last modified on Wednesday, 24 August 2011 14:18

Published in Architecture

Tagged under data center pue energy efficiency

Social sharing

RELATED ITEMS (BY TAG)

- Online Tech Expands Data Center Operations
- Data Center Consolidation: Balancing Cost-Savings vs. Business Continuity/Disaster Recovery
- And Now – A Data Center World
- Some Gaps In Energy Ratings
- The Platinum LEED Data Center From GE

Add comment

Name (required)

E-mail (required, but will not display)

Website

Notify me of follow-up comments

Refresh

Send

JComments

Back To Top