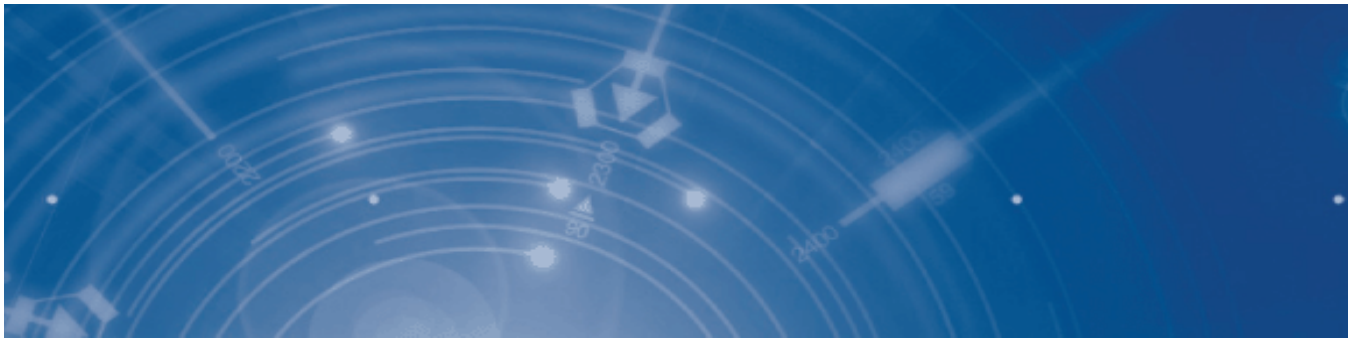

How IT Management Can "Green" the Da

22 January 2008

Rakesh Kumar, Simon Mingay

Gartner RAS Core Research Note G00153396

Data centers waste large amounts of energy. IT management can help them to "go green" by impleme
redundancies and inefficiencies in hardware and software, and staying ahead of the trends affecting pc

Overview

In data centers, large-scale inefficiencies (in areas such as distribution and power conversion) occur throughout the energy chain, from the potential of the fuel source to generation and distribution. IT management can help enterprises that run data centers improve their energy efficiency and develop a comprehensive green strategy.

Key Findings

- Data centers tend to lose massive amounts of energy, but they can be outfitted and retrofitted for energy efficiency.
- Data center servers and information and communication technology (ICT) devices often have low utilization rates and extraneous software.
- IT managers should put together an environmental strategy for data centers that includes metrics, modeling, consolidation of equipment and machines, and possible decommissioning.
- Future data centers can be greened from the ground up, with energy efficiency and other environmental considerations integrated in site selection, structure design and construction, and choice of equipment.

Recommendations

- Assess data center energy efficiency holistically, applying an end-to-end strategy that integrates the key components of the energy picture. Work closely with all the players along the energy chain.
- Establish goals, processes and responsibilities, targeting energy efficiency, as well as waste management, asset management, capacity management, support services and facilities management.
- Measure power consumption in the ICT infrastructure as granularly as possible, using

metering tools at all levels (building, hardware and facilities components) to determine the amount of power going into the data center.

Analysis

Data centers are energy-intensive, so it follows that energy consumption is one of their biggest environmental issues. However, a green data center will broaden its environmental strategy beyond energy efficiency, gleaning the maximum amount of production from the minimum amount of materials and energy, without compromising performance, resilience and security.

Such an approach requires an end-to-end integrated view that includes the configuration of the building, energy efficiency, waste management, asset management, capacity management, technology architecture, support services, energy sources and operations. This research provides IT managers with an outline of the trends affecting data centers and offers strategies with which to address these changes.

Return to Top

Architectural Changes of Core Server and Storage Technologies

More than 80% of server shipments are made up of x86 boxes; however, enterprise data center topologies vary. Most large clients maintain legacy and niche technologies, such as Alpha, Bull GCOS, Fujitsu and BS2000. Moving off these platforms is costly, because the application functionality has to be "re-created" on a new platform, which can be difficult and expensive.

In some cases, users grow their legacy infrastructures. For example, many IBM mainframe customers have grown their installed MIPS during the past four years and will continue to do so. However, most new workloads are hosted in x86 high-density rack server environments, which are powerful enough for most current application workloads and will be used increasingly for database engines.

Utilization of infrastructure remains low for most hardware platforms. A typical x86 server uses between 5% and 10% of its available capacity during a 24-hour period — reduced instruction set computer (RISC) Unix systems are slightly better, at 10% to 20%. Emerging technologies — such as server and storage virtualization, dynamic workload management, subprocessor partitioning and metering/monitoring tools — will evolve rapidly during the next few years, resulting in increasingly real-time infrastructures. Enterprises planning data center space should factor in these changes, which means users will need to maintain legacy hardware environments or they'll be forced to move to newer platforms.

Return to Top

The Energy Needed to Power and Cool ICT Infrastructures Is Likely to Increase

During the next 10 years, the increased energy appetite of processors, servers, storage devices and network appliances will increase the energy required to power and cool the ICT hardware infrastructure. A shortage of prime quality data center space (especially in Europe) to host new infrastructure, as well as the rapidly increasing cost to build new data centers, will add to the problem.

For many organizations, actual energy consumption will outpace this underlying trend for a period of time. Users will get onto a "technology platform runway," and, by sticking to it and increasing the volume of hardware technology, they'll experience a sharp increase in energy consumption. However, the industry will see innovation at many levels (including servers, management software, liquid cooling and blowers), and, at some point, they'll switch to newer technologies and almost instantly reduce their energy consumption.

Over a period of time, the curve will again go upward, and users will need another correction. This "saw tooth" type of curve will, however, always trend upward. Users should adopt a philosophical, technical and financial posture that takes advantage of this scenario.

Return to Top

Managing the Evolving Energy Picture

In a data center, massive inefficiencies occur throughout the energy chain, from the potential of the original fuel source through generation and distribution. Many of these inefficiencies lie outside the control of IT departments, resulting from how power is generated (different sorts of power stations) and distributed in countries' national electricity grids. Once the electricity reaches the data center, cooling, uninterruptible power source (UPS) and other "nonproductive" uses expend resources. This is compounded when energy gets to the server and IT devices, which often have low utilization rates and extraneous software.

Although data center energy efficiency is not about improving cooling, power management and the power source, these areas are a good place to start the greening process. The IT management team should work on driving energy efficiency across the software, technology infrastructure, architecture and design. This endeavor requires:

- Close cooperation with architects, software engineers and data center operations, as well as collaboration with the facilities and real estate teams.
- Sourcing decisions based on coefficiency, including shared service models, such as software as a service (SaaS), that have the potential to be energy- and eco-efficient. Business process and application architects and designers need to take heed, because these decisions will affect service levels, technology choices and implementation — all of which will influence energy efficiency.
- Models that provide a granular picture of energy costs, floor space and infrastructure topology of data centers and related offices. These models will be important to developing projections based on the growth of IT equipment and changes to the data center or office layout, and will enable the consideration of different scenarios of equipment deployment and internal financial management (such as chargeback). For example, a model using a computational fluid dynamics (CFD) analysis can determine how many more servers can go into a data center, where they should be located and whether any changes need to be made to an air-conditioning system.
- A systematic program for consolidating machines and workloads to use available spare capacity and maximize return on investment (ROI) in server technology. This will delay the need to purchase more of the newer, high-density hardware and could push back the manifestation of the energy problems by months, if not years.
- Procurement of new servers that run at approximately a 60% to 70% utilization rate. This will ensure that they better manage power and cooling issues at the initiation of hardware deployments. Increasing ICT equipment utilization levels from less than 20% has dramatic and multiplying benefits for a data center's energy consumption.
- Use of software tools to achieve optimal utilization rates. These tools include virtualization software from such companies as Microsoft and VMware, as well as better workload

applications. However, these tools are of little use if the organizations don't change their operational processes so they benefit from the software. For example, multiple virtual machines in a single server should ensure that production partitions can be run next to test partitions in the same box.

- A rigorous decommissioning process that physically removes equipment once it becomes redundant. The IT organization must also identify low-utilization devices and consider consolidating and decommissioning them. Examine devices that are plugged in and drawing power for no purpose.
- Disposal strategies. Organizations should comply with directives and legislation — Waste Electrical and Electronic Equipment (WEEE) and Restriction of Hazardous Substances (RoHS) — as well as go further by developing processes for disposal/break-up/smelting and the recovery of metals.
- Measurements of power consumption at the power distribution unit level to gain a higher threshold of energy for core computing equipment.

As with any transition, start by measuring key metrics to get a status report. The Green Grid and a number of others have proposed two useful data center metrics: power usage effectiveness (PUE) and data center efficiency (DCE). These measure the proportion of power consumed doing productive work vs. energy used doing nonproductive activities, such as cooling. Both metrics provide insights and can highlight areas for improvements in most data centers. This assumes that the power consumed by the IT equipment is productive. Often, it isn't: A server farm run 24/7, with 15% utilization rate during office hours, would not show up in these metrics.

Herein lines the challenge for the industry and each individual enterprise: How do you measure what comes out of the data center? At the moment, there is no objective metric of "real data center efficiency." An individual data center might be able to create one for internal purposes, depending on the nature of the work undertaken by the infrastructure, such as the total power consumed by ICT equipment/general ledger transactions. Such a measure captures an element of business activity, but it's imprecise. However, in the long term, it requires an energy management architecture.

Work with the enterprise to encourage combining heat and power (CHP). The practical and economic availability of efficient power generation and the ability to apply sourcing strategies vary significantly, depending on the technology, location, availability of grants, planning permission, the local utility providers and so on. However, CHP offers high levels of efficiency, reliability and a positive ROI in most circumstances. It also offers increased security of supply when backed up by the grid. Photovoltaics offer an expensive, partial opportunity for some, but they are a long way from being a practical and economic solution for most.

Google is running a pilot in California with a 1.6MW array, with the intention of expanding it considerably. However, for piloting and public relations reasons, it's too early for much more than that. Wind is economic and scalable, but intermittent, and it's not practical in many areas. Geothermal offers potential, but is not economical for most at the moment. Fuel cells are experimental at this stage — T-Systems is running a 250kW pilot in a Munich data center — but they do offer potential for the future. Although these are examples of emerging technologies, users should consider all options.

Point to the benefits of renewable energy. Off the grid, it's expensive and increasingly scarce as many enterprises look for quick and easy fixes, but renewable energy can be purchased for other parts of the operation where their application is less expensive.

Return to Top

Conclusion: Data Centers Tend to Lose Energy, but They Can Improve

Data centers of the future have the advantage of being "groomed green" from the ground up. Designers can choose a site based on energy security, cost and source. They can develop that site, probably on a modular basis, building the physical structure — its fixtures and fittings — to be energy-efficient and use relatively low-impact materials and construction techniques. They can outfit the center with chillers, as well as heating, ventilation and air conditioning systems, and introduce recycling and alternative resources, such as solar, wind, geothermal, hydro, fuel cell and CHP. They can monitor tools and manage server efficiencies to measure and manage the environmental footprint. Gartner's advice is to build all new large facilities with the chilled fluid plumbing at the outset.

Established data centers have no such advantages, and enterprises will have to contend with many challenges. Massive inefficiencies can occur along the energy chain, from the original fuel source to generation and distribution. Once the electricity reaches the data center, cooling, UPS and other nonproductive uses expend resources. This is compounded when energy gets to the server and to ICT devices, which often have low utilization rates and extraneous software.

IT managers can ameliorate this tendency by helping data centers to improve these conditions. For example, they could:

- Select and design support services that meet cost, service levels and performance requirements and operate in a coefficient manner.
- Select ICT equipment according to an assessment of lifetime environmental impact.
- Encourage the use of distributed (local) power generation using CHP, which can offer efficiencies and a positive ROI.
- Conduct a study of PUE to provide insights into the energy efficiency of the supporting data center infrastructure.
- Consider the energy efficiency of the workload through server utilization tests and other methods.
- Set goals for energy efficiency, waste, asset and capacity management, support services and facilities management.
- Measure power consumption in the ICT infrastructure and be as granular as possible, using submetering — at the very least, measure the power going into the data center.

Return to Top