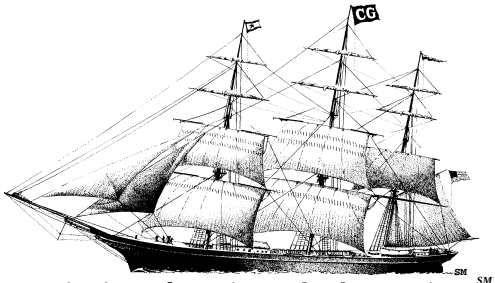
Clipper Notes™
VENDOR-AGNOSTIC EXPLANATIONS AND ADVICE FOR THE INFORMATION TECHNOLOGY BUYER



Navigating Information Technology Horizons℠

Published Since 2007 | Report #TCG2008061 | November 24, 2008

Capping Energy Demand in the Data Center — "It's Not Easy Being Green!"

Analyst: David Reine

Management Summary

The headlines in the news and the political debate on the airwaves all bring to light serious issues that are affecting our lives. The cost of energy and the fears of global warming require our immediate attention. Fortunately, each of us has some control over our personal pocketbook. We do not have to wait for the nabobs of Washington to act. We can and should help ourselves.

The most obvious way to improve the environment and reduce energy consumption is our transportation. If we move to energy efficient cars, we can reduce air pollution and gas consumption. Hybrid automobiles that use less gas are high on the list of answers; however, the cost of a new car can be prohibitive in today's economic climate, even if it saves you $50 per week in gas and improves the air that our children breathe. However, there are easier and less expensive steps that we can take in our own homes, using readily available technology. Take, for example, our room lights. Today, when you enter a room you flip a switch and turn on the light. If you replace the light bulb in each lamp with an energy-saving bulb, you can extend the life of the bulb *and* reduce the amount of energy consumed. Even though the bulb may cost more, you will use less energy and save money! Another readily available, and inexpensive, alternative is on the wall near that light switch: the thermostat. By installing a setback thermostat, you can automatically regulate the amount of energy consumed by time-of-day and day-of-the-week. Again, this could save a sizeable amount of energy and money for you. In either case, you might not notice the difference.

These same strategies also could be applied to the enterprise data center, although the savings involved in greening the environment would pale when compared to the resources being wasted in the typical data center on a daily basis. Due to the server sprawl that has spread through the data center like a virulent virus adding complexity to an already complex environment, anywhere from 75% to 85% of server resources may be wasted (unused) at any given time. Storage, too, often has a lot of waste associated with it – primarily from over-allocation and storing rarely- or never-accessed information on a too-costly device that provides a higher level of service than needed.

Along with CPU cycles and terabytes, the data center is also wasting the energy to run and cool servers with no simple means to cap the power requirement, to reduce energy use. This waste has been driving up the total cost of ownership (TCO) of the data center for years. Until the recent jump in the cost of energy, and concern about the carbon footprint of the enterprise, the data center generally has ignored it, often because the cost of electricity was in the facilities budget and not in the IT budget. However, the hypergrowth of servers and storage has driven many data centers to the out edges of floor space, cooling and energy availability. **In order to reduce the TCO of IT, the data center must get this situation under control and improve the utilization of the data center infrastructure before being forced to build a new data center**. ***When the enterprise goes green, it can save a lot of green!***

Many data centers already have implemented or have begun to implement two major strategy initiatives to improve energy utilization and lower TCO – *consolidation* and *virtualization*. Unfortunately, these only scratch the surface of the problem. As that famous PBS philosopher, *Kermit the Frog*, has said often: ***It's not easy being green!*** To learn more about reducing energy consumption in the data center, please read on.

IN THIS ISSUE

Wasted Energy in the Data Center

Do you remember that old line that would be heard frequently from the support center: *If is doesn't work, try plugging it in.* Well, today, we hear just the opposite: *If you're not using it, unplug it!* Most data centers are wasting energy all day, every day. Very few, if any, servers or storage devices operate at close to 100% efficiency. In fact, if your data center did have a totally-efficient server, you rightfully would be concerned that it did not have the resources necessary for growth. Unfortunately, the majority of installed *x86* servers are running at utilization rates of 15 to 25%[1]. Not only are CPU cycles being wasted, 75% of the energy required to run the server farm and storage arrays and cool the data center environment are also being spent without significant benefit to the enterprise. This adds significantly to data center TCO, and is especially true for any enterprise with a server population deployed prior to 2006, due to the current availability of multi-core servers based on Intel and AMD processors.

Prior to 2006, most data centers deployed mono- or dual-socket servers in a scale-out environment, with single-core CPUs. The IT staff installed these platforms with a single application for a dedicated user community. Even though many of these processors had the capability to regulate the power consumed, there were few, if any, power management tools available to measure or regulate them. **The ability to throttle the energy might have been there, but it was unmanaged**. Prior to 2006, however, there was much less concern about the amount of electricity coming through the wall. For most enterprises, it was a utility; there was always enough to satisfy their needs. That is no longer true. Today, many enterprises have to be concerned about the amount of energy available (i.e., is there more available within an existing data center or do we need to build another data center just to get more power?), not just the cost. **Before the IT department deploys another platform, either server or storage, it must determine if there is enough electricity to go around, or the entire data center could be in jeopardy of overloading or overheating.**

In elementary school, we learned about the three "R"s – *Reading*, '*Riting*, and '*Rithmetic*. Today, the IT staff must learn three *new "R"s* regarding the use of energy in the data center. They must learn to **reduce** energy consumption, **to use less**, and **reclaim** wasted energy, **to use no more**, yet have sufficient energy on-hand to support future expansion, so that the enterprise can **Rock 'n' Roll** through the years to come, and **use energy more wisely** when necessary. (OK, maybe that is four "R"s.) In order to get to where you want to be, however, you need to know where you are!

- *How much energy does your data center consume today?*
- *What does that consumption cost?*
- *How much of that cost is the result of waste due to poor utilization of both servers and storage?*
- *How much is due to over-provisioning in support of high availability?*
- *What kind of growth of server processing and storage are you experiencing?* (See Exhibit 1 on the next page.)

Allocating $50K in new capital expenditures may make a great deal of sense if it can save you $100K in recurring costs due to inefficient use of energy, floor space, and software licensing. Before you can even set a target, however, the IT staff must determine what the energy consumption level is today. Fortunately, most have a pretty good idea of how close to the edge they already are.

Surprisingly, even yesterday's technology has some features that enable you to measure and control the power level of your existing x86 servers. Intel, for example, has *Demand-Based Switching with Enhanced Intel SpeedStep* technology in their *Xeon* architecture, while AMD has *PowerNow* technology in *Opteron* that enables the CPU to lower power consumption without compromising performance, based upon CPU utilization. AMD's *Optimized Power Management* feature can reduce processor power based upon workload, improving performance/watt. **These energy-limiting features empower the IT staff with platform and software energy-management features to help lower average power consumption and heat generation while helping to maintain application performance, *but only if they are enabled*.** Thus, it is important to include energy management features in data center management software.

The Solution is in Your Hands

Today, every CIO is asking the same questions: *What can I do to increase performance while reducing TCO and to enable more processing power and storage capacity?* Plus, as part of reducing TCO, *what can I do to lower energy costs?*

The first, and most obvious, action item is to use less energy by reducing the IT infrastructure. Consolidation and virtualization of under-

[1] Typically true for older one and two core servers that are more that two years old. However, just because you have a more modern server, it doesn't mean that your actual utilization is much higher. It all depends on how focused you were on server efficiency (i.e., optimizing server utilization) when you consolidated via virtualization and also on how smart your applications are in how they handle increasing workloads.

Exhibit 1 — Three Modes of Data Center Growth and Efficient Energy Use

• **Growth is Moderate** – annual growth is less than 50% (typical of many smaller enterprises)	• **Use Less Energy** – By consolidating and virtualizing the current application set on existing servers and storage, the IT staff can reduce infrastructure, power consumption, and floor space, extending the life span of the existing data center.
• **Growth is High** – annual growth is greater than 50% to, say, 100% (typical of most larger enterprises)	• **Use No More Energy** – By replacing existing servers with new platforms using multi-core technology within the same power envelope, the IT staff can double performance capability of the existing facility and hold power consumption constant; likewise, by replacing 500GB drives with 1TB drives, the data center can double capacity at the same energy level.
• **Growth is Hyperactive** – annual growth is 200% to 1000%, or more, typical of the largest Web 2.0 firms.	• **Use More Wisely** – By combining newer platform technology with consolidation, virtualization, and data deduplication, the data center can enable growth within the same data center even with a controlled rise in power consumption.

utilized servers and storage can improve the TCO picture significantly by eliminating underperforming platforms, and the associated administrative and maintenance costs, reducing data center energy requirements, freeing up vital floor space for growth, and lowering software licensing fees by evolving to a scale-up[2] architecture. It is very simple, really! If you can minimize the infrastructure, you can eliminate many energy-sapping components, such as power supplies and fans.

Today's servers all employ multi-core processors – not only Intel's *Xeon* and AMD's *Opteron*, but IBM's *POWER6*, and Sun's *SPARC*, as well. **With the appearance of 4-, 6-, and 8-core CPUs, each new generation of processors can double, triple, or even quadruple server performance, not just within the same footprint, but usually within the same power envelope, as well, improving the all-important *performance/-watt* metric, ensuring that you use no more energy.** All of these new processors have embedded virtualization features, enabling the IT staff to improve server utilization and efficiency. In fact, some of the new CPUs have multiple threads per core further enhancing their virtualization capability, as long as the application software has the ability to take advantage of this feature. The IT staff needs to have the skills, or access to professional services, to determine whether a single-core CPU with a higher clock speed may be able to outperform a multi-core, multi-threaded processor in the execution of some mission-critical applications.

In addition to the evolution of multi-core processors in the data center, the enterprise is also undergoing a paradigm shift in the packaging of server technology. With the introduction of processor blades, the data center can share a major portion of the platform infrastructure over multiple servers, enabling them to take advantage of economies of scale to manage energy consumption even further. This is a significant development, especially in high availability environments, because the IT staff often intentionally over-provisions power and cooling functionality to ensure sufficient resources in the event of a component failure. The data center must take advantage of the embedded processor functionality to cap power consumption on a CPU, blade, or node basis, in order to reclaim energy previously wasted. By regulating the amount of power that any specific processor can consume, the IT staff can ensure that there will be sufficient resources for the on-demand scalability required in mission- and business-critical environments. In some cases, this may result in the consumption of more power in order to allow for both business continuity and growth. **If you will need to use more energy, it is important to use more energy wisely**.

Unfortunately, servers are not the only resource consuming power in the data center. While data centers have been consolidating applications on multi-core servers and bladed environments, storage arrays are rapidly becoming the largest consumers of energy in the data center, because storage is doubling in capacity every 12-to-18 months at most enterprises. New energy-efficient storage arrays are appearing, enabled with such features as thin provisioning and storage virtualization to reduce the amount of physical storage actually required in a consolidated architecture. In addition, another paradigm-shifting technology is

[2] See the issue of Clipper Notes dated September 23, 2008, entitled *Perceiving the Dark Side of the Moon – Knowing When Scale-up Computing Makes Sense*, and available at http://www.clipper.com/research/TCG2008048.pdf.

making inroads in the data center. *MAID*[3] technology enables the IT staff to control the amount of energy consumed by storage by turning off unaccessed drives or those not likely to be accessed (as determined by stated policies) or, in some cases, slowing those disk devices in the Tier-2 array that are not used very frequently. In addition, just as we have seen in the server arena with multi-core CPUs, the data center can consolidate storage on larger capacity disk devices, currently up to 1TB, although migrating to 1TB drives could imply the requirement for RAID 6, increasing overhead for the additional parity drive needed by that technology. As with the CPUs, however, doubling the capacity of your existing disk drives with more energy-efficient drives will also lower your energy consumption. Consolidating storage will also achieve a similar benefit as seen with servers: the data center will be able to reclaim valuable floor space. Freeing the data center from over-crowding will reduce power consumption, improve the air flow, and reduce cooling costs, as well as providing much-needed space for growth.

Data Deduplication is another hot technology to cool the data center. Implementing data dedupe within your storage environment enables the IT staff to put the same amount of backup data in less space, in some cases 80% to 90% less space, although based upon the content of the data you are saving and the frequency of those backups, your mileage may vary! You will need to weigh the acquisition cost of the data deduplication environment against your potential savings in the storage architecture.

You also need to examine closely the construction of your Tier-2 storage framework. Many enterprises have transitioned recently to a D2D environment for their backup and archiving needs. The IT staff needs to re-evaluate the advantages, and costs, of tape. The implementation of tape for long-term backup and archiving requirements can have significant TCO benefits for the enterprise. In fact, in terms of energy, disk has 290 times the cost of tape for long-term backup[4]. Returning to a tape environment for long-term storage can also have significant side benefits in helping you lower the TCO of the data center. Tape technologies, such as *LTO*, are removable from the system to help prevent sabotage, protect critical data from disaster at offsite locations, and enable the data center with a highly scalable architecture for future growth. With a current compressed capacity of 1.6TB per LTO-4 cartridge, LTO has a roadmap leading to compressed capacity of 6.4TB for LTO-6, with performance growth from a current 240MB/sec to a projected throughput for LTO-6 of 540MB/sec. In addition, with WORM technology included, the cartridges are "audit-safe," and with encryption built-in, LTO cartridges are safe to recertify and recycle, as any attempt to retrieve information from them would be futile.

Conclusion

Helping to save the planet always makes ecological sense. When saving energy reduces the TCO of the data center environment, then you are also making a fiscal statement as well. That is a win-win situation if I ever saw one. **That is exactly what creating a green environment within your data center does: it enables the enterprise to put more green on the bottom line.**

It is essential for the enterprise to know what its carbon footprint is. To do that, the data center must put in place energy-efficient platforms, both server and storage, and the means to measure and manage energy consumption. In the coming months, there will be Energy Star guidelines to assist the concerned CIO in procuring and deploying a "green" environment[5]. You must include these ratings in your procurement guidelines to ensure compliance.

Today's technology is in place to assist the enterprise in reducing its carbon footprint:

- Multi-core processors for mission- and business-critical servers that maintain a consistent thermal envelope with power capping to reduce energy consumption;
- TB disk drives to consolidate complex islands of storage into shared pools of information;
- Multi-TB tape drives to reduce the amount of energy, and space, required to store data for years, and, in fact, decades.

It is up to us to engage these technologies in our workplaces to help control the environment and, oh yes, help to reduce the TCO of the data center.

[3] Massive Array of Idle Disks

[4] See the issue of Clipper Notes dated February 13, 2008, entitled *Disk and Tape Square Off Again – Tape Remains King of the Hill with LTO-4*, and available at http://www.clipper.com/research/TCG2008009.pdf.

[5] At this time, the EPA has excluded blades from Energy Star consideration.

About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

David Reine **is Director, Enterprise Systems for The Clipper Group.** Mr. Reine specializes in enterprise servers, storage, and software, strategic business solutions, and trends in open systems architectures. He joined The Clipper Group after three decades in server and storage product marketing and program management for Groupe Bull, Zenith Data Systems, and Honeywell Information Systems. Mr. Reine earned a Bachelor of Arts degree from Tufts University, and an MBA from Northeastern University.

- ***Reach David Reine via e-mail at dave.reine@clipper.com or at 781-235-0085 Ext. 123. (Please dial "123" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group *Captain's Log***, **The Clipper Group Voyager**, **Clipper Notes**, and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.