

How green is your storage?

BY LARRY FREEMAN, CRAIG NUNES,
PHIL TSIHLIS, AND AL THOMASON

Here are 10 technologies to improve storage efficiency and reduce power consumption.

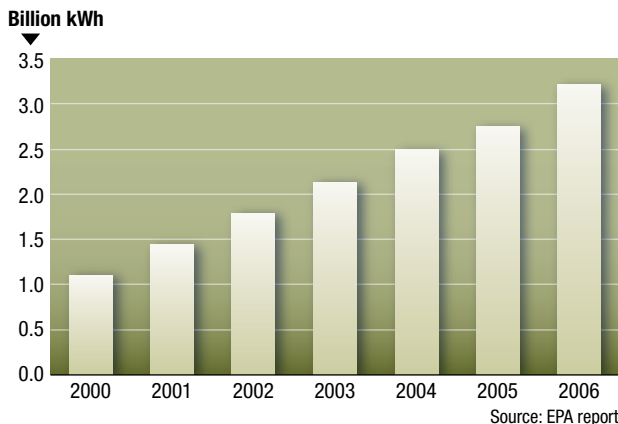
An Environmental Protection Agency (EPA) report to Congress [1] compared the energy consumption of four primary data-center components: high-end servers, midrange servers, networking equipment, and storage devices (see Table). In this study, data storage devices had the highest power consumption growth rate (191%) and the highest overall power consumption (3.2 billion kWh.)

Power consumption (billion kWh)

| | 2000 | % of total | 2006 | % of total | Growth |
|----------------------|------------|-------------|------------|-------------|------------|
| Data storage devices | 1.1 | 18% | 3.2 | 32.3% | 191% |
| Network equipment | 1.4 | 23% | 3 | 30.3% | 114% |
| Midrange servers | 2.5 | 41% | 2.2 | 22.2% | -12% |
| High-end servers | 1.1 | 18% | 1.5 | 15.2% | 36% |
| Total | 6.1 | 100% | 9.9 | 100% | 62% |

According to the report, power consumption of data storage devices maintained a steady growth rate during the period (see figure, below). Left unchecked, this growth rate will soon encumber the power requirements of other data-center components.

Data storage power consumption



Adding to the problem of rising power requirements in the data center is the fact that every watt of power consumed by IT equipment requires at least another watt for infrastructure, which includes cooling, UPS, lighting, and losses through power distribution. In other words, *each watt saved in the data center is two watts earned!*

In light of this, data storage vendors have been optimizing power efficiency through various design aspects of their products. The

benefits of this are the following:

- Reduced storage power requirement “balances” the overall power consumption of IT equipment in the data center (i.e., provides more available power to the servers); and
- Reduced storage power requirements decrease the overall data-center power requirements, reducing operational costs.

“Green storage” is a simple way to describe data storage (or storage networking) products that can be configured for optimal energy efficiency and power savings. However, the components that constitute green storage, and the techniques for making storage “greener,” are still largely unknown or misunderstood.

This article summarizes several approaches to reducing storage power consumption, including high-efficiency power supplies, high-capacity disk drives, and often-overlooked space-saving software options.

High-efficiency power supplies

A large amount of data-center power is lost due to poorly designed power supplies with low efficiency ratings. According to recent studies, inefficient power supplies in data-center equipment contribute a power loss of 50% or more during periods of low power consumption.

Designing products with an efficient power profile involves two steps. First, power supply-rated output specifications should be closely matched to the components that are being provided this power. Second, power supplies should deliver an optimal amount of power efficiency across the entire product load range. Poorly designed IT products using overrated power supplies that continually operate within their lowest efficiency load range needlessly drain power from the data center.

Today, high-efficiency power supplies are available in disk and tape systems, fabric switches and directors, and other storage network appliances. Deploying products designed for energy efficiency incrementally reduces the overall data-center power bill.

High-capacity disk drives

The latest storage systems use disk drives with the highest capacities in history. Using these high-capacity drives allows users to drive down watts/terabyte in the data center. For example, migrating data stored

[1] EPA - Report to Congress on Server and Data Center Energy Efficiency (www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf)

SNIA's Green Storage Initiative

SNIA's Green Storage Initiative (GSI) comprises 23 storage system and component manufacturers representing all categories of storage solutions. The initiative is committed to:

- Educating the user and vendor communities about the importance of power conservation and environmental sustainability in shared storage environments;
- Formulating and publicizing best practices for energy-efficient storage networking; and
- Promoting storage-centric applications that reduce the storage footprint and associated power requirements.

For more information, go to www.snia.org/forums/green/.

on legacy 36/73/146GB Fibre Channel drives to newer, higher-capacity Fibre Channel or Serial Attached SCSI (SAS) drives can significantly improve power/cooling profiles. Similarly, migrating infrequently accessed application data to high-capacity SATA tiers will substantially improve storage energy efficiency.

In high-performance applications, one drawback of high-capacity drives is reduced I/O throughput. "Wide striping" overcomes this by allowing high-performance applications to be spread across many (tens or even hundreds) more disks. Because wide striping allows many volumes to share a given drive, utilization is much higher. Therefore, application data can sustain a high number of IOPS with high-capacity drives, avoiding the necessity of low-capacity, high-rpm, energy-intensive drives.

Advanced RAID techniques

When high-capacity disk drives are used for storage devices, larger amounts of data are stored per drive. Therefore, care must be taken to ensure data reliability is not compromised. In the past, this protection was commonly addressed through RAID-1 mirroring. Today, space-efficient RAID implementations have become more commonplace, including single-parity RAID 5 and recent RAID-6 innovations such as dual parity and P+Q algorithms. When compared to data mirroring, these technologies offer up to 70% greater storage utilization, resulting in fewer power-consuming

drives needed to provide protection against drive failures.

Thin provisioning

A key problem faced by storage administrators is storage quota allocation. How much physical storage space should be assigned for each application? Knowing that an overflowing data volume has many unpleasant side effects, administrators commonly overprovision their disk quotas. If they think an application will require a single terabyte, he might decide to allocate 2TB, "just in case," to accommodate for growth, or to adjust for a miscalculation of the storage space actually consumed by the application.

But what if the application does not grow as expected, or if the miscalculation was on the short side? The result is wasted space—space that cannot be used by any other application. By some estimates, 60% or more of disk storage remains unused simply because of this type of over provisioning. Unused disk capacity, however, continues to draw power and contributes to the overall data-center electricity bill.

The problem of over provisioning can be solved through thin provisioning, where administrators can create "flexible" volumes that appear to the application to be a certain size but in reality are much smaller physically. Thin-provisioning technology provides substantial improvements in storage sizing. Data volumes can be resized quickly and dynamically as application requirements change.

The bottom-line impact of thin provisioning is a reduction in physically allocated storage, and direct savings in data-center power, heat, and cooling requirements.

Data de-duplication

The average disk volume contains thousands or even millions of duplicate data objects. As data objects are created, modified, distributed, backed up, and archived, duplicate data quickly begins to proliferate throughout the organization. The result is inefficient use of storage resources. Data de-duplication helps to prevent this inefficiency.

Typically, data de-duplication divides stored data objects into smaller blocks. Each block of data has a digital "signature," which is compared to all other signatures in the data volume. If an exact block match exists, then the

duplicate block is discarded and its disk space is reclaimed. De-duplication can be implemented across a wide variety of applications and file types, including primary data, backup data, and archival data. By implementing de-duplication, users can reclaim up to 95% of their storage space.

Note that combining thin provisioning and data de-duplication has an additive effect on the efficiency of storage. De-duplicated volumes are sometimes oversized when the de-duplication savings ratio proves to be greater than predicted. De-duplicated volumes are also sometimes oversized intentionally to account for some amount of growth. Thin provisioning eliminates this additional capacity overhead pre-allocated for de-duplication.

Writable snapshots

Storage administrators must often allocate substantial storage space for enterprise test operations, such as application release rollouts and bug fix testing. In addition, organizations that rely on large-scale simulations for comprehensive testing, analysis, and modeling can incur large costs associated with providing additional storage space for these tests.

In the past, to address this issue, administrators would simply make complete copies of a data set as their "test set." By offering writable snapshots, vendors provide application "clone" functionality where application copies can be created as temporary, writable copies. Furthermore, these copies can be created instantly, with minimal storage requirements.

This is accomplished by creating a writable "snapshot" of the primary dataset and storing only the data changes between a parent volume and a clone. All unchanged data remains on primary storage and is used by both the primary application and the secondary clone copy. Multiple snapshot copies can be created from a single primary dataset, enabling users to perform multiple test and development simulations and compare the characteristics of each dataset after the testing is complete.

Data compression

Used for decades in tape drives and home computers, data compression has recently appeared in data centers in two specific areas:

- External data compression appliances that compress data "on-the-fly" as data is stored on storage systems; and

- Disk-to-disk (D2D) backup devices, such as virtual tape libraries (VTLs), which use data compression to reduce the amount of storage required by backup copies.

These appliances are generally based on the Lempel-Ziv compression algorithm and can offer 50% or greater storage savings.

Flash drives

Solid-state flash drives use flash memory to store and access data. Because there are no mechanical components in flash drives, they provide faster response times and consume 38% less energy on average vs. traditional mechanical disk drives, resulting in a significant power consumption reduction in a transaction-per-second comparison. When deployed in combination with hard disk drives, flash drives provide an ultra-high-performance “tier” of storage for transaction application environments requiring optimal performance, while leveraging hard drive-based tiers for less demanding applications. Solid-state flash drives offer the ability to achieve high performance without sacrificing energy costs.

Standby and spin-down modes

Just as tape media uses no energy when it is not being accessed, if one is able to spin down unused or underutilized disk drives, noticeable power savings can be seen. Technologies such as MAID (massive array of idle disks) are now available, and potential future developments in intelligent controllers will allow disk drives to enter a series of reduced power states. Although spinning

down a disk drive has a positive benefit on energy consumption, there is a likewise obvious impact to data retrieval response times.

Another technology advocated by some vendors is standby mode for the entire storage system. The idea is that during off-peak hours, disk controllers that are not being accessed could go into “sleep” mode to save even more energy. This is similar to modes currently used by PCs—microprocessors in most storage systems have the identical capability. A standby mode invoked, say, between midnight and 6 AM would represent a 25% daily power savings.

Virtualization

By virtualizing servers, several “guest” servers can operate on a single physical server, reducing the overall number of servers in the data center and their associated power consumption. Virtualization technologies can also be applied to disk-based storage systems to reduce the amount of physical storage needed, and hence reduce the overall power consumption. Though in many ways thin provisioning provides virtualization, it can also extend beyond this technology.

By abstracting storage elements, the administrator is able to allocate physical resources that match the current usage needs—associating a virtual resource to high-performance storage, or more energy-efficient storage. Besides allowing for dynamic changes in virtual as well as physical volume sizes, virtualization can allow the transparent migration of application data between different classes of

storage. For example, a project might initially be deployed on high-performance Fibre Channel drives, then as the project finishes its peak usage and moves more into a maintenance phase, the data can be transparently migrated to a more energy-efficient storage subsystem to take advantage of the better watts/gigabyte ratio of higher-capacity drives.

In this way, overall energy needs can be reduced, but more importantly can be appropriately assigned to the correct power/performance storage transparently to the application. This transparent placement of data can be extended with copy services mentioned above to allow administrators to custom-fit an application’s needs to the available storage resources.

Summary

Energy consumption continues to be one of the most significant portions of the cost of operating a data center. Finding ways to increase energy efficiency is of critical importance to data-center managers and has become a significant public policy issue. Using the technologies described in this article, you can take significant steps toward reducing data-storage power consumption, leading to a “greener” data center. ©

This article was written on behalf of the Storage Networking Industry Association’s Green Storage Initiative. The authors’ company affiliations are: Larry Freeman (NetApp), Craig Nunes (3PAR), Phil Tsihlis (EMC), and Al Thomason (IBM).



NetApp™
Go further, faster