SNIA
Green Storage Initiative

Best Practices for Energy Efficient Storage Operations Version 1.0

October 2008

Authors:
Tom Clark, Brocade – GSI Chair
Dr. Alan Yoder, NetApp – GSI Governing Board

SNIA

Introduction

The energy required to support data center IT operations is becoming a central concern worldwide. For some data centers, additional energy supply is simply not available, either due to finite power generation capacity in certain regions or the inability of the power distribution grid to accommodate more lines. Even if energy is available, it comes at an ever increasing cost. With current pricing, the cost of powering IT equipment is often higher than the original cost of the equipment itself. The increasing scarcity and higher cost of energy, however, is being accompanied by a sustained growth of applications and data. Simply throwing more hardware assets at the problem is no longer viable. More hardware means more energy consumption, more heat generation and increasing load on the data center cooling system. Companies are therefore now seeking ways to accommodate data growth while reducing their overall power profile. This is a difficult challenge.

Data center energy efficiency solutions span the spectrum from more efficient rack placement and alternative cooling methods to server and storage virtualization technologies. The SNIA's Green Storage Initiative was formed to identify and promote energy efficiency solutions specifically relating to data storage. This document is the first iteration of the SNIA GSI's recommendations for maximizing utilization of data center storage assets while reducing overall power consumption. We plan to expand and update the content over time to include new energy-related storage technologies as well as SNIA-generated metrics for evaluating energy efficiency in storage product selection.

Some Fundamental Considerations

Reducing energy consumption is both an economic and a social imperative. While data centers represent only ~2% of total energy consumption in the US, the dollar figure is approximately $4B annually. In terms of power generation, data centers in the US require the equivalent of six 1000 MegaWatt power plants to sustain current operations. Global power consumption for data centers is more than twice the US figures. The inability of the power generation and delivery infrastructure to accommodate the growth in continued demand, however, means that most data centers will be facing power restrictions in the coming years. Gartner predicts that by 2009, half of the world's data centers will not have sufficient power to support their applications.[1] An Emerson Power survey projects that 96% of all data centers will not have sufficient power by 2011.[2] Even if there was a national campaign to build alternative energy generation capability, new systems would not be online soon enough to prevent a widespread energy deficit. This simply highlights the importance of finding new ways to leverage technology to increase energy efficiency within the data center and accomplish more IT processing with fewer energy resources.

In addition to the pending scarcity and increased cost of energy to power IT operations, data center managers face a continued explosion in data growth. Since 2000, the amount of corporate data generated worldwide has grown from 5 exabytes (5 billion gigabytes) to over 300 exabytes, with projections of ~1 zetabyte (1000 exabytes) by 2010.[3] This data must be stored somewhere. The sustained growth of data requires new tools for data management, storage allocation, data retention and data redundancy.

The conflict between the available supply of energy to power IT operations and the increasing demand imposed by data growth is further exacerbated by the operational requirement for high availability access to applications and data. Mission-critical applications in particular are high energy consumers and require more powerful processors, redundant servers for failover, redundant networking connectivity, redundant fabric pathing, and redundant data storage in the form of mirroring and data replication for disaster recovery. These top tier applications are so essential for business operations, however, that the doubling of server and storage hardware elements and the accompanying doubling of energy draw have been largely unavoidable. Here too, though, new green storage technologies and best practices can assist in retaining high availability of applications and data while reducing total energy requirements.

Shades of Green

The quandary for data center managers is in identifying which new technologies will actually have a sustainable impact for increasing energy efficiency and which are only transient patches whose initial energy benefit quickly dissipates as data center

[1] "Gartner Says 50 Percent of Data Centers Will Have Insufficient Power and Cooling Capacity by 2008," Gartner Inc. Press Release, November 29, 2006

[2] "Emerson Network Power Presents Industry Survey Results That Project 96 Percent of Today`s Data Centers Will Run Out of Capacity by 2011" Emerson Press Release, November 16, 2006

[3] "A Forecast of Worldwide Information Growth Through 2010", IDC White Paper, March, 2007

requirements change. Unfortunately, the standard market dynamic that eventually separates weak products from viable ones has not had sufficient time to eliminate the green pretenders. Consequently, analysts often complain about the 'greenwashing' of vendor marketing campaigns and the opportunistic attempt to portray marginally useful solutions as the cure to all the IT manager's energy ills.

Within the broader green environmental movement greenwashing is also known as being 'lite green' or sometimes 'light green'. There are, however, other shades of green. Dark green refers to environmental solutions that rely on across-the-board reductions in energy and material consumption. For a data center, a dark green tactic would be to simply reduce the number of applications and associated hardware and halt the expansion of data growth. Simply cutting back, however, is not feasible for today's business operations. To remain competitive, businesses must be able to accommodate growth and expansion of operations.

Consequently, viable energy efficiency for ongoing data center operations must be based on solutions that are able to leverage state-of-the-art technologies to do much more with much less. This aligns to yet another shade of environmental green known as "bright green". Bright green solutions reject both the superficial lite green and the Luddite dark green approaches to the environment and rely instead on technical innovation to provide sustainable productivity and growth while steadily driving down energy consumption. The following SNIA GSI best practices include many bright green solutions that accomplish the goal of energy reduction while increasing productivity of IT storage operations.

Although the Best Practices recommendations listed below are numbered sequentially, no prioritization is implied. Every data center operation has different characteristics and what is suitable for one application environment may not work in another.

These recommendations collectively fall into the category of "silver buckshot" in addressing data center storage issues. There is no single silver bullet to dramatically reduce IT energy consumption and cost. Instead, multiple energy efficient technologies can be deployed in concert to reduce the overall energy footprint and bring costs under control. Thin provisioning and data de-duplication, for example, are distinctly different technologies that together can help reduce the amount of storage capacity required to support applications and thus the amount of energy-consuming hardware in the data center. When evaluating specific solutions, then, it is useful to imagine how they will work in concert with other products to achieve greater efficiencies.

Best Practice #1 – Manage Your Data

A significant component of the exponential growth of data is the growth of redundant copies of data. By some industry estimates, over half of the total volume of a typical company's data exists in the form of redundant copies dispersed across multiple storage systems and client workstations. Consider the impact, for example, of emailing a 4MB PowerPoint attachment to 100 users instead of simply sending a link to the file. The corporate email servers now have an additional 400 MB of capacity devoted to redundant copies of the same data. Even if individual users copy the attachment to their local drives, the original email and attachment may languish on the email server for months before the user tidies their Inbox. In addition, some users may copy the attachment to their individual share on a data center file server, further compounding the duplication. And to make matters worse, the lack of data retention policies can result in duplicate copies of data being maintained and backed up indefinitely.

This phenomenon is replicated daily across companies of every size worldwide, resulting in ever increasing requirements for storage, longer backup windows and higher energy costs. A corporate policy for data management, redundancy and retention is therefore an essential first step in managing data growth and getting storage energy costs under control. Many companies lack data management policies or effective means to enforce them because they are already overwhelmed with the consequences of prior data avalanches. Responding reactively to the problem, however, typically results in the spontaneous acquisition of more storage capacity, longer backup cycles and more energy consumption. To proactively deal with data growth, begin with an audit of your existing applications and data and begin prioritizing data in terms of its business value.

Although tools are available to help identify and reduce data redundancy throughout the network, the primary outcome of a data audit should be to change corporate behavior. Are data sets periodically reviewed to ensure that only information that is relevant to business is retained? Does your company have a data retention policy and mechanisms to enforce it? Are you educating your users on the importance of managing their data and deleting non-essential or redundant copies of files? Are your Service Level Agreements (SLAs) structured to reward more efficient data management by individual departments? Given that data generators (i.e., end users) typically do not understand where their data resides or what resources are required to support it, creating policies for data management and retention can be a useful means to educate end users about the consequences of excessive data redundancy.

SNIA

Proactively managing data also requires aligning specific applications and their data to the appropriate class of storage. Without a logical prioritization of applications in terms of business value, all applications and data receive the same high level of service. Most applications, however, are not truly mission-critical and do not require the more expensive storage infrastructure needed for high availability and performance. In addition, even high-value data does not typically sustain its value over time. As we will see in the recommendations below, aligning applications and data to the appropriate storage tier and migrating data from one tier to another as its value changes can reduce both the cost of storage and the cost of energy to drive it. This is especially true when SLAs are structured to require fewer backup copies as data value declines.

Best Practice #2 – Select the Appropriate Storage RAID Level

Storage networking provides multiple levels of data protection, ranging from simple CRC checks on data frames to more sophisticated data recovery mechanisms such as RAID. RAID guards against catastrophic loss of data when disk drives fail by creating redundant copies of data or providing parity reconstruction of data onto spare disks.

RAID 1 mirroring creates a duplicate copy of disk data, but at the expense of doubling the number of disk drives and consequently doubling the power consumption of the storage infrastructure. The primary advantage of RAID 1 is that it can withstand the failure of one or all of the disks in one mirror of a given RAID set. For some mission-critical environments, the extra cost and power usage characteristic of RAID 1 may be unavoidable. Accessibility to data is sometimes so essential for business operations that the ability to quickly switch from primary storage to its mirror without any RAID reconstruct penalty is an absolute business requirement. Likewise, asynchronous and synchronous data replication provide redundant copies of disk data for high availability access and are widely deployed as insurance against system or site failure. As shown in Best Practices #1, however, not all data is mission critical and even high value data may decrease in value over time. It is therefore essential to determine what applications and data are absolutely required for continuous business operations and thus merit more expensive and less energy efficient RAID protection.

RAID 5's distributed parity algorithm enables a RAID set to withstand the loss of a single disk drive in a RAID set. In that respect, it offers the basic data protection against disk failure that RAID 1 provides, but only against a single disk failure and with no immediate failover to a mirrored array. While the RAID set does remain online, a failed disk must be reconstructed from the distributed parity on the surviving drives in the set, possibly impacting performance. Unlike RAID 1, however, RAID 5 only requires one

SNIA

spare drive in a RAID set. Fewer redundant drives means less energy consumption as well as better utilization of raw capacity.

By adding two additional drives, RAID 6 can withstand the loss of two disk drives in a RAID set, providing a higher availability than RAID 5. Both solutions, however, are more energy efficient than RAID 1 mirroring (or RAID 1+0 mirroring and striping) and should be considered for applications that do not require an immediate failover to a secondary array.

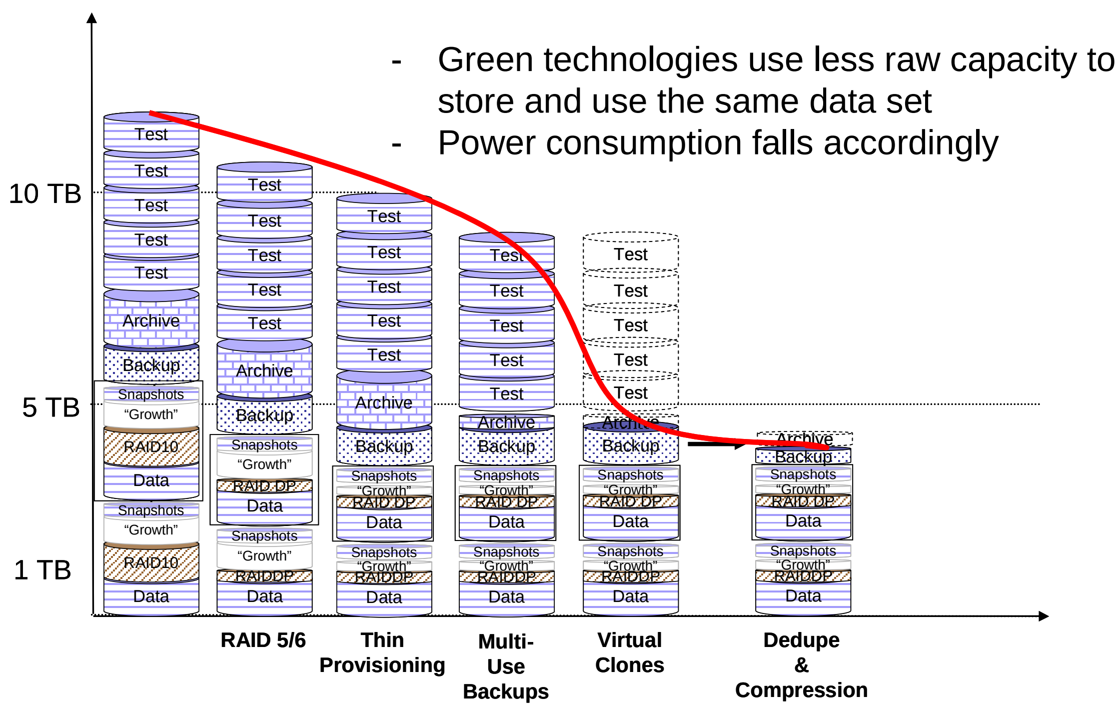

Figure 1
Software Technologies for Green Storage

Alan Yoder - NetApp

As shown in Figure 1, the selection of the appropriate RAID levels to retain high availability data access while reducing the storage hardware footprint can enable incremental green benefits when combined with other technologies.

Best Practice #3 – Leverage Storage Virtualization

Storage virtualization refers to a suite of technologies that create a logical abstraction layer above the physical storage layer. Instead of managing individual physical storage arrays, for example, virtualization enables administrators to manage multiple storage systems as a single logical pool of capacity, as shown in Figure 2.

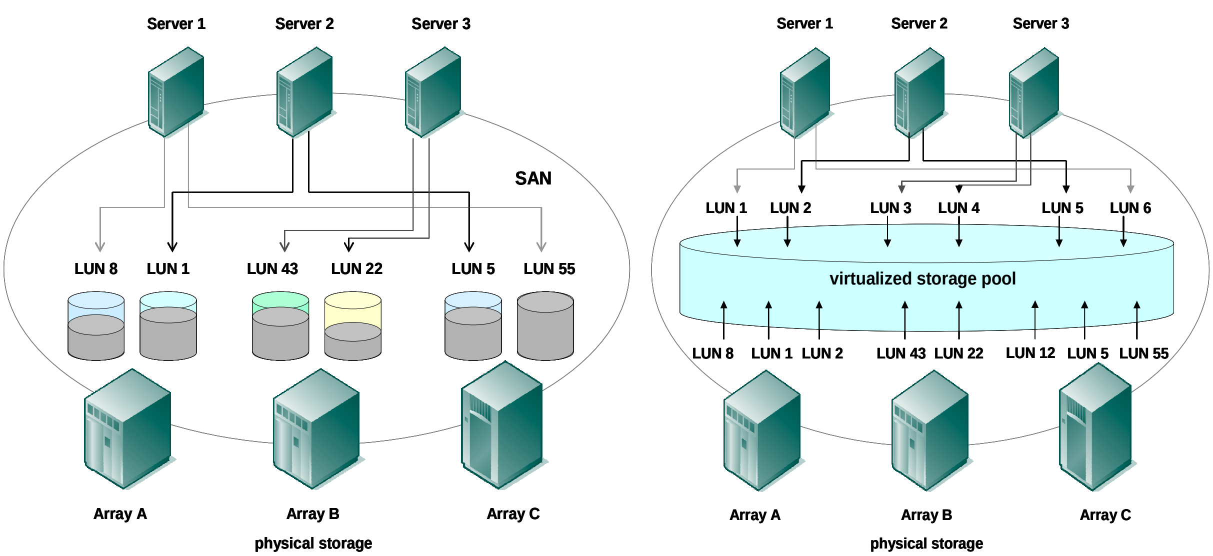

Figure 2
Storage Virtualization: Technologies for Simplifying Data Storage and Management
T. Clark – Addison-Wesley - used with permission from the author

On its own, storage virtualization is not inherently more energy efficient than conventional storage management but can be used to maximize efficient capacity utilization and thus slow the growth of hardware acquisition. By combining dispersed capacity into a single logical pool, it is now possible to allocate additional storage to resource-starved applications without having to deploy new energy-consuming hardware. Storage virtualization is also an enabling foundation technology for thin provisioning, resizeable volumes, snapshots and other solutions that contribute to more energy efficient storage operations.

Best Practice #4 – Use Data Compression

Compression has long been used in data communications to minimize the number of bits sent along a transmission link and in some storage technologies to reduce the amount of data that must be stored. Depending on implementation, compression can impose a performance penalty because the data must be encoded when written and decoded (decompressed) when read. Simply minimizing redundant or recurring bit patterns via compression, however, can reduce the amount of processed data that is stored by one half or more and thus reduce the amount of total storage capacity and hardware required.

Not all data is compressible, though, and some data formats have already undergone compression at the application layer. JPEG, MPEG and MP3 file formats, for example, are already compressed and will not benefit from further compression algorithms when written to disk or tape.

When used in combination with security mechanisms such as data encryption, compression must be executed in the proper sequence. Data should be compressed before encryption on writes and decrypted before decompression on reads.

Best Practice #5 – Incorporate Data Deduplication

While data compression works at the bit level, conventional data deduplication works at the disk block level. Redundant data blocks are identified and referenced to a single identical data block via pointers so that the redundant blocks do not have to be maintained intact for backup (virtual to disk or actual to tape). Multiple copies of a document, for example, may only have minor changes in different areas of the document while the remaining material in the copies have identical content. Data deduplication also works at the block level to reduce redundancy of identical files. By retaining only unique data blocks and providing pointers for the duplicates, data deduplication can reduce storage requirements by up to 20:1. As with data compression, the data deduplication engine must reverse the process when data is read so that the proper blocks are supplied to the read request.

Data deduplication may be done either in band, as data is transmitted to the storage medium, or in place, on existing stored data. In band techniques have the obvious advantage that multiple copies of data never get made, and therefore never have to be hunted down and removed. In place techniques, however, are required to address the immense volume of already stored data that data center managers must deal with.

Best Practice #6 – File Deduplication

File deduplication operates at the file system level to reduce redundant copies of identical files. Similar to block level data deduplication, the redundant copies must be identified and then referenced via pointers to a single file source. Unlike block level data deduplication, however, file deduplication lacks the granularity to prevent redundancy of file content. If two files are 99% identical in content, both copies must be stored in their entirety. File deduplication therefore only provides a 3 or 4 to 1 reduction in data volume in general. Rich targets such as full network-based backup of laptops may do much better than this, however.

Best Practice #7 – Thin Provisioning of Storage to Servers

In classic server-storage configurations, servers are allocated storage capacity based on the anticipated requirements of the applications they support. Because exceeding that storage capacity over time would result in an application failure, administrators typically over-provision storage to servers. The result of fat provisioning is higher cost, both for the extra storage capacity itself and in the energy required to support additional spinning disks that are not actively used for IT processing.

Thin provisioning is a means to satisfy the application server's expectation of a certain volume size while actually allocating less physical capacity on the storage array or virtualized storage pool. This eliminates the under-utilization issues typical of most applications, provides storage on demand and reduces the total disk capacity required for operations. Fewer disks equate to lower energy consumption and cost and by monitoring storage usage the storage administrator can add capacity only as required.

Best Practice #8 – Leverage Resizeable Volumes

Another approach to increasing capacity utilization and thus reducing the overall disk storage footprint is to implement variable size volumes. Typically, storage volumes are of a fixed size, configured by the administrator and assigned to specific servers. Dynamic volumes, by contrast, can expand or contract depending on the amount of data generated by an application. Resizeable volumes require support from the host operating system and relevant applications, but can increase efficient capacity utilization to 70% or more. From a green perspective, more efficient use of existing disk capacity means fewer hardware resources over time and a much better energy profile.

Best Practice #9 – Writeable Snapshots

Application development and testing are integral components of data center operations and can require significant increases in storage capacity to perform simulations and modeling against real data. Instead of allocating additional storage space for complete copies of live data, snapshot technology can be used to create temporary copies for testing. A snapshot of the active, primary data is supplemented by writing only the data changes incurred by testing. This minimizes the amount of storage space required for testing while allowing the active non-test applications to continue unimpeded.

Best Practice #10 – Deploy Tiered Storage

Storage systems are typically categorized by their performance, availability and capacity characteristics. Formerly, most application data was stored on a single class of storage system until it was eventually retired to tape for preservation. Today, however, it is possible to migrate data from one class of storage array to another as the business value and accessibility requirements of that data changes over time. Tiered storage is a combination of different classes of storage systems and data migration tools that enables administrators to align the value of data to the value of the storage container in which it resides. Because second-tier storage systems typically use slower spinning or less expensive disk drives and have fewer high availability features, they consume less energy compared to first-tier systems. In addition, some larger storage arrays enable customers to deploy both high-performance and moderate-performance disks sets in the same chassis, thus enabling an in-chassis data migration.

A tiered storage strategy can help reduce your overall energy consumption while still making less frequently accessed data available to applications at a lower cost per gigabyte of storage. In addition, tiered storage is a reinforcing mechanism for data retention policies as data is migrated from one tier to another and then eventually preserved via tape or simply deleted.

Best Practice #11 – Solid State Storage

Solid state storage still commands a price premium compared to mechanical disk storage, but has excellent performance characteristics and much lower energy consumption compared to spinning media. While solid state storage may not be an option for some data center budgets, it should be considered for applications requiring high performance and for tiered storage architectures as a top-tier container.

Best Practice #12 – MAID and Slow-Spin Disk Technology

High performance applications typically require continuous access to storage and thus assume that all disk sets are spinning at full speed and ready to read or write data. For occasional or random access to data, however, the response time may not be as critical. MAID (massive array of idle disks) technology uses a combination of cache memory and idle disks to service requests, only spinning up disks as required. Once no further requests for data in a specific disk set are made, the drives are once again spun down to idle mode. Because each disk drive represents a power draw, MAID provides inherent green benefits. As MAID systems are more accessed more frequently, however, the energy profile begins to approach those of conventional storage arrays.

Another approach is to put disk drives into slow spin mode when no requests are pending. Because slower spinning disks require less power, the energy efficiency of slow spin arrays is inversely proportional to their frequency of access.

Occasionally lengthy access times are inherent to MAID technology, so it is only useful when data access times of several seconds—the length of time it takes a disk to spin up—can be tolerated.

Best Practice #13 – Tape Subsystems

As a storage technology, tape is the clear leader in energy efficiency. Once data is written to tape for preservation, the power bill is essentially zero. Unfortunately, however, businesses today cannot simply use tape as their primary storage without inciting a revolution among end users and bringing applications to their knees. Although the obituary for tape technology has been written multiple times over the past decade, tape endures as a viable archive media. From a green standpoint, tape is still the best option for long term data retention.

Best Practice #14 – Fabric Design

Fabrics provide the interconnect between servers and storage systems. For larger data centers, fabrics can be quite extensive with thousands of ports in a single configuration. Because each switch or director in the fabric contributes to the data center power bill, designing an efficient fabric should include the energy and cooling impact as well as rational distribution of ports to service the storage network.

A mesh design, for example, typically incorporates multiple switches connected by interswitch links (ISLs) for redundant pathing. Multiple (sometimes 30 or more) meshed switches represent multiple energy consumers in the data center. Consequently, consolidating the fabric into higher port count and more energy efficient director chassis and core-edge design can help simplify the fabric design and potentially lower the overall energy impact of the fabric interconnect.

Best Practice #15 – File System Virtualization

By some industry estimates, 75% of corporate data resides outside of the data center, dispersed in remote offices and regional centers. This presents a number of issues, including the inability to comply with regulatory requirements for data security and backup, duplication of server and storage resources across the enterprise, management and maintenance of geographically distributed systems and increased energy consumption for corporate-wide IT assets. File system virtualization includes several technologies for centralizing and consolidating remote file data, incorporating that data into data center best practices for security and backup and maintaining local response-time to remote users. From a green perspective, reducing dispersed energy inefficiencies via consolidation helps lower the overall IT energy footprint.

SNIA

Best Practice #16 – Server, Fabric and Storage Virtualization

Data center virtualization leverage virtualization of servers, the fabric and storage to create a more flexible and efficient IT ecosystem. Server virtualization essentially deduplicates processing hardware by enabling a single hardware platform to replace up to 20 platforms. Server virtualization also facilitates mobility of applications so that the proper processing power can be applied to specific applications on demand. Fabric virtualization enables mobility and more efficient utilization of interconnect assets by providing policy-based data flows from servers to storage. Applications that require first class handling are given a higher quality of service delivery while less demanding application data flows are serviced by less expensive paths. In addition, technologies such as NPIV (N_Port ID Virtualization) reduce the number of switches required to support virtual server connections and emerging technologies such as FCoE (Fibre Channel over Ethernet) can reduce the number of hardware interfaces required to support both storage and messaging traffic. Finally, storage virtualization supplies the enabling foundation technology for more efficient capacity utilization,

snapshots, resizeable volumes and other green storage solutions. By extending virtualization end-to-end in the data center, IT can accomplish more with fewer hardware assets and help reduce data center energy consumption.

File system virtualization can also be used as a means of implementing tiered storage with transparent impact to users through use of a global name space.

Best Practice #17 – Flywheel UPS Technology

Flywheel UPSs, while more expensive up front, are several percent more efficient (typically > 97%), easier to maintain, more reliable and do not have the large environmental footprint that conventional battery-backed UPSs do. Forward-looking data center managers are increasingly finding that this technology is less expensive in multiple dimensions over the lifetime of the equipment.

Best Practice #18 – Data Center Air Conditioning Improvements

The combined use of economizers and hot-aisle/cold aisle technology can result in PUEs of as low as 1.25. As the PUE (Power Usage Effectiveness ratio) of a traditional data center is often over 2.25, this difference can represent literally millions of dollars a year in energy savings.

Economizers work by using outside air instead of recirculated air when doing so uses less energy. Obviously climate is a major factor in how effective this strategy is: heat and high humidity both reduce its effectiveness.

There are various strategies for hot/cold air containment. All depend on placing rows of racks front to front and back to back. As almost all data center equipment is designed to draw cooled air in the front and eject heated air out the back, this results in concentrating the areas where heat evacuation and cool air supply are located.

One strategy is to isolate only the cold aisles and to run the rest of the room at hot aisle temperatures. As hot aisle temperatures are typically in the 95° F range, this has the advantage that little to no insulation is needed in the building skin, and in cooler climates, some cooling is gotten via ordinary thermal dissipation through the building skin.

Another strategy is to isolate both hot and cold aisles. This reduces the volume of air that must be conditioned, and has the advantage that humans will find the building temperature to be more pleasant.

In general, hot aisle/cold aisle technologies avoid raised floor configurations, as pumping cool air upward requires extra energy.

Best Practice #19 – Increased Data Center temperatures

Increasing data center temperatures can save significant amounts of energy. Ability to do this is dependent in much part on excellent temperature and power monitoring capabilities, and on conditioned air containment strategies. Typical enterprise class disk drives are rated to 55° C (131° F), but disk lifetime suffers somewhat at these higher temperatures, and most data center managers think it unwise to get very close to that upper limit. Even tightly designed cold aisle containment measures may have 10 to 15 degree variations in temperature from top to bottom of a rack; the total possible variation plus the maximum measured heat gain across the rack must be subtracted from the maximum tolerated temperature to get a maximum allowable cold aisle temperature. So the more precisely that air delivery can be controlled and measured, the higher the temperature one can run in the "cold" aisles.

Benefits of higher temperatures include raised chiller water temperatures and efficiency, reduced fan speed, noise and power draw, and increased ability to use outside air for cooling through an economizer.

Best Practice #20 – Work with Your Regional Utilities

Some electrical utility companies and state agencies are partnering with customers by providing financial incentives for deploying more energy efficient technologies. If you are planning a new data center or consolidating an existing one, incentive programs can provide guidance for the types of technologies and architectures that will give the best results.

What the SNIA is Doing About Data Center Energy Usage

The SNIA Green Storage Initiative is conducting a multi-pronged approach for advancing energy efficient storage networking solutions, including advocacy, promotion of standard metrics, education, development of energy best practices and alliances with other industry energy organizations such as The Green Grid. Currently, over 20 SNIA members have joined the SNIA GSI as voting members.

A key requirement for customers is the ability to audit their current energy consumption and to take practical steps to minimize energy use. The task of developing metrics for measuring the energy efficiency of storage network elements is being performed by the SNIA Green Storage Technical Work Group (TWG). The SNIA GSI is supporting the technical work of the GS-TWG by funding laboratory testing required for metrics development, formulation of a common taxonomy for classes of storage and promoting GS-TWG metrics for industry standardization.

The SNIA encourages all storage networking vendors, channels, technologists and end users to actively participate in the green storage initiative and help discover additional ways to minimize the impact of IT storage operations on power consumption. If, as industry analysts forecast, adequate power for many data centers will simply not be available, we all have a vital interest in reducing our collective power requirements and make our technology do far more with far less environmental impact.

For more information about the SNIA Green Storage Initiative, link to:

http://www.snia.org/forums/green/

To view the SNIA GSI Green Tutorials, link to:

http://www.snia.org/education/tutorials#green

Appendix A

Green Storage Terminology

Active Power
The energy consumption of a system when powered on and under normal workload.

Blanking Plate
Typically a solid metal plate that mounts over unused data center rack slots to maintain efficient air flow through the rack.

Bright Green
Applying new technologies to enhance energy efficiency while maintaining or improving productivity.

Dark Green
Addressing energy consumption by the acoss-the-board reduction of energy consuming activities.

DAS
Direct-attached storage. One or more disk drives directly connected to a server or host, typically via parallel SCSI cabling, Serial-Attached SCSI (SAS) or Serial ATA (SATA).

Data Center Virtualization
The application of various virtualization technologies at the server, fabric and storage layers to create a more flexible end-to-end operational environment.

Data Compression
Bit-level reduction of redundant bit patterns in a data stream via encoding. Typically used for WAN transmissions and archival storage of data to tape.

Data Deduplication
Block-level reduction of redundant data by replacing duplicate data blocks with pointers to a single good block.

Economizer
Heat exchanger technology typically used to leverage colder external air to provide data center cooling. Dry side economizers use cooler outdoor air; wet side economizers use cooling towers or chillers.

Exabyte
One billion gigabytes.

File Deduplication
Reduction of file copies by replacing duplicates with pointers to a single original file.

Fly Wheel Technology
Use of momentum of a spinning disk or wheel to temporarily generate electricity in the event of a data center power failure. Fly wheel technology provides the bridge between normal power distribution and backup diesel generators and can replace conventional battery rooms.

Gbps/W
Gigabits per second per Watt. A metric for evaluating the fabric bandwidth provided per fixed unit of energy.

GB/W
Gigabytes per Watt. A metric for evaluating the storage capacity provided per fixed unit of energy.

Greenwashing
A byproduct of excessive marketing and ineffective engineering.

GSI
Green Storage Initiative. Within the SNIA, initiatives are composed of interested member companies for conducting marketing, education, promotion and development of specific storage technologies. The SNIA GSI is promoting green storage technologies and supports the technical work of the GS TWG.

Hot Aisle / Cold Aisle
Arranging IT equipment in racks so that heat is exhausted in designated aisles while cool air is supplied in the alternating aisles.

Idle Power
The power consumption of a system when powered on but with no active workload.

IOPs/W
Input/Output operations per second per Watt. A metric for evaluating storage I/O performance per fixed unit of energy.

Lite (or Light) Green
Solutions or products that purport to be energy efficient but which only have negligible green benefits.

MAID
Massive Array of Idle Disks. A storage array that only spins up disks to active state when data in a disk set is accessed or written.

Metric
A standard unit of measurement, typically part of a system of measurements to quantify a process or event within a given domain. GB/W and IOPs/W are examples of proposed metrics that can be applied for evaluating the energy efficiency of storage systems.

MaxTTD
Maximum time to data. For a given category of storage, the maximum time allowed to service a data read or write.

NAS
Network-attached storage. An optimized platform that provides file-based access to storage using NFS or CIFS protocols over an IP network infrastructure.

Near Online Storage
Storage systems with longer maximum time to data access, typical of MAID and fixed content storage (CAS)

Non-Removable Media Library
A virtual tape backup system with spinning disks and shorter maximum time to data access compared to conventional tape.

Online Storage
Storage systems with fast data access, typical of most data center storage arrays in production environments.

PDU
Power Distribution Unit. A system that distributes electrical power, typically stepping down the higher input voltage to voltages required by end equipment. A PDU can also be a single-inlet/multi-outlet device within a rack cabinet.

RAID
Redundant Array of Independent Disks. A means of storing block data across multiple disks so that data can be recovered in the event of disk failure. RAID 1 creates a mirror of data in disk or disk set by duplicating all data writes to a secondary disk or disk set. RAID 5 calculates and writes block parity across a RAID set so data can be reconstructed from parity information in the event of a single disk failure. RAID 6 employs a dual parity scheme to enable data reconstruction in the event that two disks fail concurrently.

Removable Media Library
A tape or optical backup system with removable cartridges or disks and >80ms maximum time to data access.

Resizeable Volumes
Variable length volumes that can expand or contract depending on the data storage requirements of an application.

Raised Floor
Typical of older data center architecture, a raised floor provides space for cable runs between equipment racks and cold air flow for equipment cooling.

SAN
Storage area network. A dedicated network connecting servers to storage, typically based on Fibre Channel or iSCSI protocols.

Server Virtualization
Software that enable a single server hardware platform to support multiple concurrent instances of an operating system and applications.

Solid State Storage
A storage device based on flash or other static memory technology that emulates conventional spinning disk media.

SPOF
Single point of failure.

Storage Taxonomy
A hierarchical categorization of storage networking products based on capacity, availability, port count and other attributes. A storage taxonomy is required for the development of energy efficiency metrics so that products in a similar class can be evaluated.

Storage Virtualization
A suite of technologies that create a logical abstraction of underlying physical storage assets. Storage capacity from multiple storage arrays is pooled into a single logical resource to facilitate management of aggregate capacity instead of individual storage systems.

Thin Provisioning
Allocating less physical storage to an application than is indicated by the virtual volume size.

Tiered Storage
Deploying multiple classes of storage to align the business value of data with the value of its storage container. Data migrates from higher tiers to lower as the business value of data declines over time.

SNIA

TWG
Technical Working Group. Within the SNIA, technical work for standards and protocol development is conducted by volunteer technologists from member companies. The TWGs report to the SNIA Technical Council, which in turn reports to the SNIA Board. Green storage technical work is being conducted by the Green Storage TWG.

(Un)PluggedFest
Periodic tests conducted by the SNIA GS TWG to develop measurement methodologies for storage products, frequently held at the SNIA Technology Center in Colorado Springs

UPS
Uninterruptible power supply.

VTL
Virtual tape library. A disk-based storage system that emulates a tape library and typically provides <80ms maximum time to data access.

Writeable Snapshots
An alternative to creating a complete copy of storage data, writeable snapshot technology monitors changes to individual data blocks in the original.

Zetabyte
One thousand exabytes, or one trillion gigabytes.

About the SNIA

The Storage Networking Industry Association (SNIA) is a not-for-profit global organization, made up of some 400 member companies and 7,000 individuals spanning virtually the entire storage industry. SNIA's mission is to lead the storage industry worldwide in developing and promoting standards, technologies, and educational services to empower organizations in the management of information. To this end, the SNIA is uniquely committed to delivering standards, education, and services that will propel open storage networking solutions into the broader market. For additional information, visit the SNIA web site at www.snia.org.