IBM XIV® Storage System

Power Consumption Reinvented

White Paper

August 2008

info@xivstorage.com
www.xivstorage.com

Contents

Introduction

Reducing power consumption is one of the most significant challenges of IT centers today.

The statistics are staggering. Reports show that, in 2006, U.S. data centers were responsible for 1.5% of the nation's power consumption. Companies are reporting 10-20% annual growth in their power usage — and the numbers are growing. By the end of 2008, storage equipment is expected to lead as the biggest power consumer in data centers, surpassing servers.

With storage needs on a steep rise, it is increasingly clear that new approaches must be explored. Simply putting in more and more machines of the same kind is nearing the point of diminishing returns, with the new requirements for power, cooling, and space exceeding what these systems can provide.

Have We All Gone Green?

Not necessarily. You no longer have to be environmentally conscious to adopt green technologies. There are plenty of other compelling reasons to do so:

- You're running out of budget. With every new system you install you have to account for additional power to the system itself, as well as additional cooling. Your cost of ownership, even before factoring in training and system administration, is becoming more than your budget can accommodate.
- You simply cannot get more power. After growing accustomed to assuming that enough power is there as long as we have the means to pay for it, this fact is no longer true. More and more companies are facing a reality in which their local power company is unable to deliver any more power to their facilities.
- Your government won't allow it. In many countries, especially in densely populated areas that are home to many IT-centric businesses, it is the governments that are imposing restrictions on power usage and the ability to keep growing it year after year.

Your first priority is meeting your organization's business needs. Yet, there is a new tension in which IT decision-makers are also expected to address power efficiency. The industry is demanding solutions that truly require less power to operate and emit less heat.

The need to consider power efficiency is a key part of today's data center reality.

Apples to Apples

Comparing the power consumption of different systems is not a straightforward matter of comparing the average numbers per raw TB that vendors publish. A more meaningful comparison is to look at power consumption based on a system's ability to meet particular business requirements.

To do so, we must first define power consumption in the context of different functionalities, including a system's ability to:

- Provide a certain effective net capacity
- Sustain a certain rate of transactions
- Maintain a certain number of backup copies
- Offer a certain level of reliability

It is important to remember that performance and reliability are the outcome of a system's architecture. Systems that deliver high performance and reliability are often heavy power consumers. In fact, a smart architecture can make a difference.

Making Storage Efficient

Making your storage solution power-efficient involves two basic precepts:

- Use less power for each terabyte
- Use fewer terabytes

This plain strategy requires that a storage system offer two key attributes:

- **Low power-consumption components —** The overall consumption of a system starts with its building blocks.
- **Optimal use of all components —** The ability to maximize use of the current system before expanding it or buying a new one.

The IBM XIV® Storage System is architected such that resources never go unused. It is able to sustain and even exceed regular tier-1 performance and reliability levels despite its use of power-efficient SATA drives as its building blocks.

Building Blocks Count

A system is only as efficient as its components. For the overall system to consume less power, it must be built from the most power-efficient building blocks available.

Today's storage systems are built with 3.5" disk drives that differ in capacity, spin rate and, consequently, power efficiency. In evaluating the comparative power efficiency of SATA and FC drives, SATA drives are the clear winner. And here's why:

Most of the power to a disk is used by its moving parts: spinning the plates and moving the head. The lower spin rate of SATA drives translates to a 25-30% reduction in power consumption compared to FC drives. Furthermore, the substantially increased capacity of today's SATA drives — 1 TB compared to 450 GB in FC drives, at best, and more commonly 146 GB — means less power is used to drive a lot more capacity. Comparing the disks' efficiency in terms of watt per raw TB, we find that SATA drives are three to fifteen times more efficient. This gap will only widen when 2 TB drives are introduced in the coming year.

However, everything has its price. The reason FC drives are still used in most primary storage systems today is that they provide higher performance. An FC drive with 15000 RPM compared with a SATA drive with 7200 RPM, for example, is able to offer half the seek time on each individual disk . A system built from small capacity FC disks provides more spindles per capacity, delivering,a higher level of performance.

This is where the perfect load balancing capability of the IBM XIV Storage System comes into play. A typical XIV system partitions and stripes each volume over all its 180 disks and, at all times, maintains equilibrium across system resources. Performance is gained by using many more spindles than you would in a traditional system, especially if you consider real life implementations over time, in which restriping of volumes is unrealistic. Furthermore, distributed caching mechanisms that span equally across the system's modules, along with high capacity internal switching, ensure that no single disk or module ever becomes the bottleneck. This revolutionary architecture is the reason that using SATA drives as the building blocks for the XIV system does not compromise performance. The system is able to offer primary storage performance for all volumes and will continue to do so over time, accommodating any future volume allocations and changes in overall capacity.

Clearly, when it comes to power efficiency, SATA drives are markedly superior. A supporting architecture that eliminates the need to compromise on performance makes SATA drives the clear-cut choice, even in the most demanding environments.

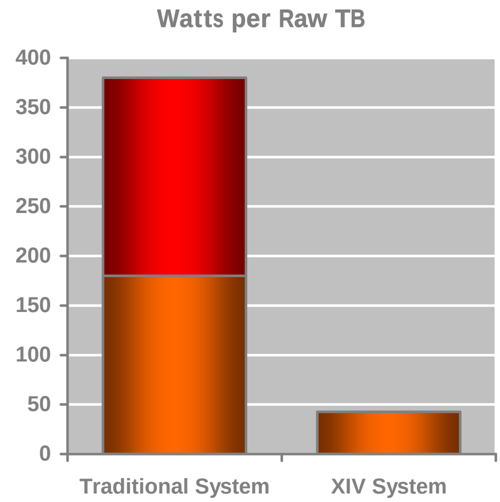
System Power Consumption

Disks are not the only component in a storage system that consumes power. A valid comparison must look at the system as a whole.

The XIV system typically consumes 7.7 kW per rack, which holds 180 TB raw capacity. This translates to consumption levels of 43 watts per raw TB (and these are expected to drop by half when 2 TB disks are used).

A typical tier-1 system comparable to the XIV system is equipped with 146 GB 15000 RPM FC drives and consumes, on average, 180–380 watts per raw TB.

The XIV system offers four to nine times improved power consumption, at comparable performance and reliability levels.



Using Less, Getting More

Using systems that are efficient per raw TB is not enough. Reducing power consumption also means maximizing use of what you have.

While this may seem trivial, analyzing typical tier-1 storage solutions reveals how, in fact, most of a solution's capacity is not put to actual use. The amount of terabytes actually used by the applications attached to the average storage system is much less than the system's net capacity.

The effective rate of a storage system is the outcome of several system attributes, all of which help enhance the system's reliability and efficiency:

- **Reliability-related attributes.** Data redundancy and hot spares make the system more reliable, at the expense of much of its raw capacity. Factoring in these attributes yields the system's "net capacity," a number that vendors publish to convey a system's actual capacity.
- **Efficiency-related attributes.** Architectural choices and system features can "stretch" each net TB. Factoring in these attributes with regard to the system's net capacity yields the system's "effective capacity" – the amount of TB that can actually be used. This is, therefore, the more relevant capacity by which to measure and compare systems.

Reliability Attributes

All storage systems must provide high reliability to ensure that data will not be lost even when components fail. To achieve this, they must use some form of data redundancy and pre-allocate spare space that, upon drive failure, allow the system to heal and go back to redundancy. Data reliability comes at the expense of a significant portion of the system's raw capacity. The actual size of that portion is mandated by the choice of RAID and the amount of spare capacity.

Choice of RAID

The choice of RAID scheme is driven by several factors, such as the necessary reliability level and the performance impact the system can endure during normal operation and during the rebuild process that takes place upon disk failure.

The RAID scheme has an immediate impact on the effective capacity rate:

- RAID-1 or mirroring is the least efficient, with 50% of raw capacity reserved for mirroring.
- RAID-5 efficiency depends on the size of the participating group. In a typical system in which a single parity disk protects six data disks, capacity utilization is 86%.
- RAID-6 with two parity disks protecting five data disks results in 71% utilization.

It is important to note that, in many cases, performance and reliability considerations associated with the choice of RAID overrule the decrease in the effective capacity rate. In fact, the majority of business critical storage systems use mirroring, despite its obvious disadvantage in terms of capacity utilization, since it requires less overhead for write operations and rebuilds upon failure are not as detrimental to the system performance as with RAID-5 and RAID-6.

The XIV system, designed with the most demanding applications in mind, uses mirroring for its data redundancy.

Allocation of Spares

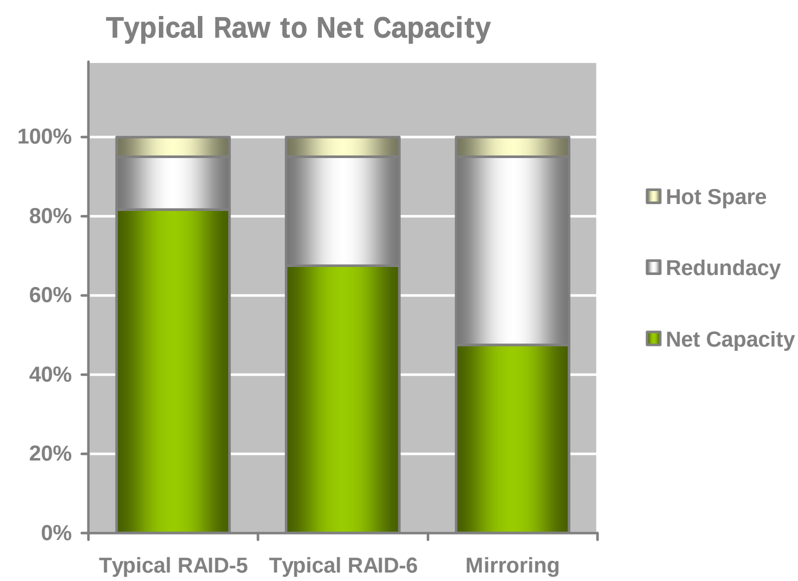
The amount of allocated spares differs from one system to another as the architecture dictates different points of failure.

Traditional systems allocate either a single spare disk per shelf, which typically holds 15 disks total. Other systems can allocate spare disks globally, reserving 3 to 12 spare disks, depending on the total number of disks in the particular system configuration. Such systems are therefore designed to use 93-98% of their raw capacity.

The IBM XIV Storage System maintains spares in the form of aggregate capacity and not as actual spare idle components. The spare capacity in each rack is allocated to match a failure of one module plus three individual disks, i.e. the equivalent capacity of 15 disks out of 180. It therefore utilizes up to 92% of its raw capacity.

Note that while similar in terms of effective rate, this approach ensures equal utilization of all 180 disks, making additional spindles available to sustain performance levels and ensure shorter rebuild times.

Overall, we find that the reliability attributes of the system limit the net capacity of a system to 46%-84% depending, for the most part, on the RAID configuration.



Efficiency Attributes

Once the net capacity of the system is established, it is up to the system's architecture and management to keep it efficient. This is where storage systems differ greatly, with three key attributes contributing to a system's effective capacity rate:

- Use of thin vs. thick provisioning
- Use of differential vs. full backup copies
- Elimination of orphaned space

The XIV system offers tremendous efficiency improvements on all three counts.

Thin Provisioning

Thin provisioning means you can define virtual volumes of high capacity and map them to much less physical storage. The storage administrator can monitor the actual use and increase the mapped capacity over time, while the applications remain indifferent to this background process, always assuming that the entire virtual volume is available to them.

In systems that do not offer thin provisioning, all volume capacity must be allocated from the first day of implementation. The administrator is obligated to predict any future needs and accommodate them all from Day One. This approach translates to the acquisition and installation of significant excessive capacity that goes unused for a long time – often months and years. Moreover, having to buy the entire predicted capacity upfront means you are stuck with yesterday's technology, unable to take advantage of the advances in storage disk technology that are likely to occur over time.

The actual savings generated by thin provisioning depends on various factors, among them the applications' data growth rate and your company's implementation standards, such as how often new systems are introduced and how much free space must be available at any given time. Overall, we estimate that, measured over three years, a thin-provisioned system will require 20-50% less capacity on average than a similar system without thin provisioning.

Differential vs. Full Backups

The ability to create periodic backups of entire volumes is essential to any enterprise storage system. Backups are required to maintain regulatory compliance, access old overwritten data, and restore corrupted data upon failure or human error.

Full backups require allocation of the entire volume size each time a backup copy is created. Differential backups consume storage space over time, requiring the allocation of space only when new data is written after the backup has been created. Differential backups thus reduce dramatically the capacity needed for each backup copy.

There is no single guideline as to the number or frequency of backups required for an application. Furthermore, while the size of a full copy is always the same, the size of a differential backup depends on the degree to which the application is write-intensive. In actual implementations, we estimate that the use of differential backups rather than full ones produces a gain of 15-30% in overall system capacity.

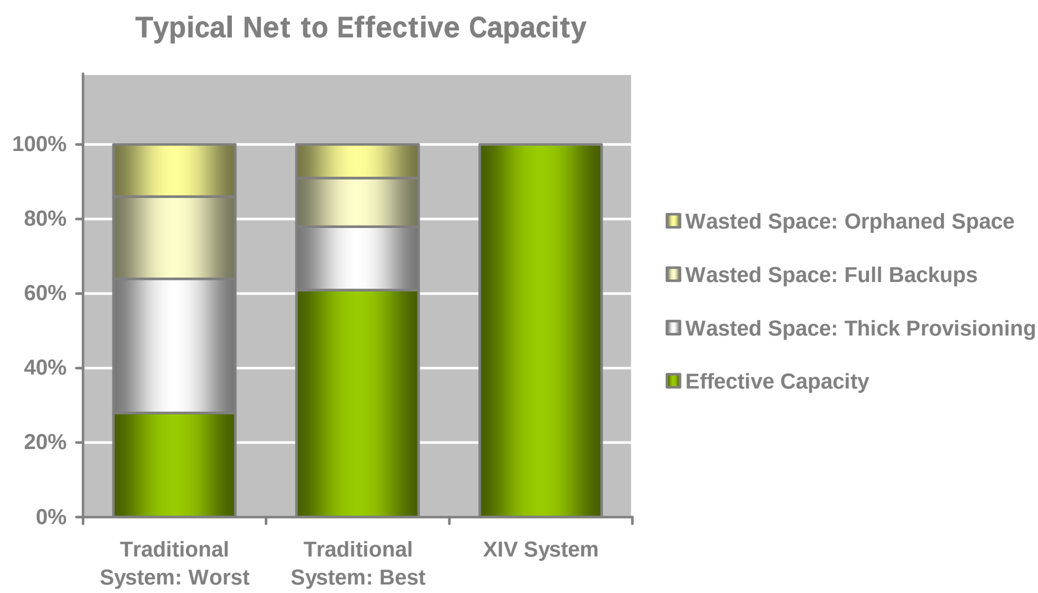
Orphaned Space

One of the faults of enterprise storage systems is that, over time, some of their capacity is effectively lost. The complexity of volume management, the need to stripe volumes to gain performance, and the everchanging nature of applications and their storage needs in an enterprise environment – all these factors eventually lead to the presence of idle storage chunks. These chunks can be reclaimed only through tedious overall system restructuring, a process that is often more expensive to the organization than buying additional storage.

The IBM XIV Storage System architecture automates the storage allocation process. The administrator need not be bothered with striping, migrating, or reallocating resources. The system is constantly kept at a balance that maximizes the performance of all the volumes and utilizes all the disks evenly. Defining new volumes, resizing existing volumes, adding more capacity, and even phasing out old hardware — all are handled internally, with no administrative effort. As a result, no space is ever lost.

It is estimated that a typical storage system has 10-20% of orphaned space that will never be reclaimed.

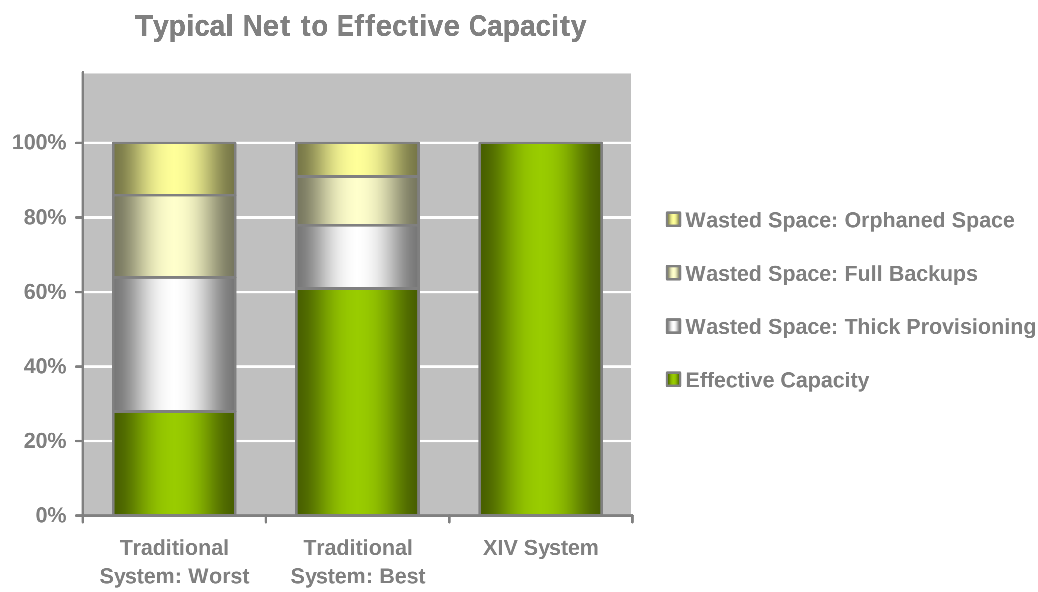
Overall, we see that, by virtue of its built-in efficiency, the XIV system uses 100% of its net capacity, compared with an estimated 28-61% net capacity used by comparable systems.



The price of compromised net capacity is clear: A system must always power 100% of its capacity yet, in traditional systems, 39-72% of this power is wasted on unneeded or unused storage space.

Conclusion and Summary

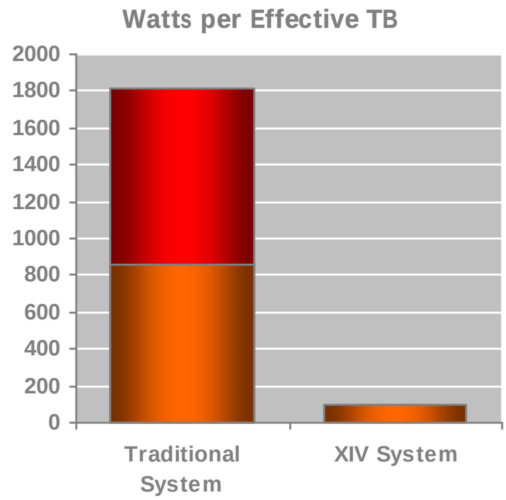
The combined effect of the reliability and efficiency attributes is such that, on average, a traditional storage system using mirroring effectively uses less than 21% of its raw capacity (37% when using RAID-5). An XIV system uses approximately 46% of its raw capacity.



In addition, the XIV system is far superior to traditional systems in terms of power consumption per raw TB. The table below summarizes a comparison of power consumption per effective TB:

System	Power / Raw TB	Effective Rate	Power / Effective TB
Traditional storage system	180-380 W	21%	857-1810 W
IBM XIV Storage System	43 W	46%	93 W

The XIV system consumes just 11-24% of the power per raw TB that would the equivalent traditional system, and a mere 5-11% of the power used to drive each effective TB.



These fantastic reductions are the outcome of the system attributes below:

Attribute	Traditional Systems	IBM XIV Storage System	Power Gain in XIV
Disk Drives	FC 146-450 GB	SATA 1 TB	Uses 4-9 times less power to drive each raw TB
Provisioning	Thick	Thin	Avoids loss of 20-50% of the net capacity by enabling over-provisioning from first day of implementation. Thin provisioning enables gradual capacity growth over time.
Backups	Full	Differential	Avoids loss of 15-30% of the net capacity that results from full backups. Differential backups only use capacity for data that has changed.
Volume Allocation	Manual: Orphaned space exists	Smart: No orphaned space	Avoids loss of 10-20% of the net capacity that results from system use over time and orphaned space that cannot be reclaimed. Smart volume allocation ensures no space is ever lost.

The XIV system uses highly efficient SATA drives, together with a smart architecture that is able to use a lot more of the net capacity, and maintains performance and reliability levels equivalent to systems built with faster, but less efficient, drives. The XIV system is a clear winner for anyone looking for a storage solution with reduced power consumption.

