Hitachi Data Systems Storage: Improving Storage Efficiency Can Be a Catalyst for IT Innovation

By Mark Peters, Senior Analyst; Jason Buffington, Senior Analyst; and Monya Keane, Research Analyst

June 2013

This ESG White Paper was commissioned by Hitachi Data Systems (HDS) and is distributed under license from ESG.
Contents

An Introduction to Storage [In]Efficiency.................................................................3
  Addressing a Need.................................................................................................3
  The Time Is Right...............................................................................................4

It’s Time to Stop Doing Storage the Old Way.....................................................5
  A Picture of Inefficiency.....................................................................................5

Meeting the Challenges—What Efficient Storage Looks Like......................8
  Dynamic Tiered Storage and Capacity Efficiency .............................................8
  Judicious Use of Solid-state Storage.................................................................9
  Data-reduction Technologies Drive Efficiency: Deduplication and Thin Provisioning .................................................................9
  Managing Archival and Cold Data ...................................................................10
  Storage Virtualization Increases Management Efficiency...........................10
  Reducing Operating Expenses: Managing TCO and Shifting Focus from CAPEX to OPEX .................................................................10
  Energy Efficiency Management......................................................................11
  Effective Management, Replication, and Storage of Multiple Copies............11
  Additional Observations About the HDS Approach to Storage Efficiency.........12

The Bigger Truth......................................................................................................13
An Introduction to Storage [In]Efficiency

At this point, the IT industry has been around for a long time. But it is still maturing and evolving in surprisingly rapid ways. Storage provides a case in point that these days, IT needs to manage storage differently.

This storage management need arises because a “business transformation” is taking place across multiple dimensions, mostly pertaining to major changes in consumption models. IT organizations (and the organizations they support) need sufficient time to:

- Address accelerated time-to-market requirements.
- Provision applications.
- Deliver the next storage service.
- Gain insight from analytics.
- Reallocate resources.

If IT can’t establish best practices for fundamental operations such as storage efficiency, how are they going to address those changing business demands identified above?

Doing what is necessary to keep operations up and running (i.e., being effective) is always IT’s top priority. When IT tries to accomplish the job with suboptimal storage infrastructures (i.e., being inefficient), the task becomes costly, operationally difficult, and even potentially overwhelming from a management perspective. On the other hand, the more efficient the storage is, the easier it becomes for IT to deliver and manage a high-availability data center.

In essence, storage efficiency equates to an ability to store and manage data in ways that consume the fewest possible resources, with little to no negative impact on performance, while lowering operational costs.

Addressing a Need

Storage inefficiency can lead to a number of undesirable and often painful issues. Consider, for example, the simple wastefulness of poorly utilized resources. Storage arrays often operate below their optimum utilization: Sixty percent to 70% might be reasonable, but it is not uncommon to find utilizations as low as 20%. The result: Unused capacity is locked in one subsystem while another array runs out of capacity. It’s akin to pouring money down a drain; we just don’t see it that way because it is “the way things have always been.” (Remarkably often in the storage realm, inefficiency has been the only practical way to achieve effectiveness. A good example: short-stroking to achieve necessary IOPS.)

Then there is the matter of silos of storage dedicated (inefficiently) to a single purpose, person, or application. Yet another matter centers on the lower levels of availability that inefficient storage often causes. When a problem occurs within an inefficient storage infrastructure, it can take longer to resolve. With downtime costs exceeding $5 million per hour in some industries (such as those that run trading environments) and constant pressures to control storage expenses occurring across all industries, the rewards for being more efficient are immense and growing daily.

The various inefficiencies described above should not be read as some ominous warning that senior managers might hear of them and demand an investigation into what data and applications their IT group could move to the cloud—after all, in most organizations, justifiable concerns exist about moving key workloads off-premises. The descriptions should instead be read as a reminder that considerable storage efficiencies are now available and proven (which was simply not true a few years ago), and IT teams can actually achieve “hero status” by implementing them as fully as possible.

After all, the opportunity cost of deploying less-than-optimally-efficient storage systems can be enormous and can easily outweigh the cost of deploying efficient storage systems in the first place.
The Time Is Right

Conditions are ripe to focus on efficiency because, again:

- Flat or declining budgets have greatly affected IT projects.
- Consumption models are changing.
- IT is trying harder than ever to meet its SLAs and thus help get the company’s products to market faster, provision IT applications faster, deliver the next storage service faster, gain insight from analytics faster, and reallocate resources faster.

Even as those budgets recover somewhat, the trumpet-call to “do more with less” blares as loudly as ever in the ears of CxOs, LOB managers, and IT managers across the globe. (Of course, a focus on efficiency is equally important in boom times to maximize profit and make the best use of staff time and resources while achieving IT objectives.)\(^1\) In other words, due to ever-tightening budgets, changing consumption models, and the need for more business agility to address all those “need-to-accelerate” challenges, we now have to do what we should have been seeking to do all along: become more efficient with the resources that we have.

The other reality making storage efficiency so important at present is the unrelenting increase of data stemming from natural application growth and from the new workloads being generated by social media; web 2.0 applications; and the creation of video, audio, photos, and similar content. A smartphone is in everyone’s pocket. A tablet computer (business as much as personal) is in many people’s laps. The ability to create and consume content requires nothing more than the press of a button. Websites and barcode readers collect more data each second—data that organizations slice and dice to identify what their customers need, or more accurately, what their customers will buy. Big data is everywhere, and the rampant copying of data sets for analytics is only one reason for it.

Other data-growth culprits include snapshots and remote replication to increase uptime and availability, and programs/initiatives to improve data protection and regulatory compliance. Those are good things, of course, but they are not always friends of efficiency.

The most common way to address the growth problem has been to toss storage capacity at it:

- You want copies for testing and development? Here’s a server and some storage.
- You’d like offsite replication? We’ll build another infrastructure stack.
- You need backup? We’ll build another.
- You need an application server? We’ll carve out a VM for you, and somehow we’ll find the storage to provision for your virtual machine image and the data it is going to need.

That strategy results in ever-expanding, unsharable silos of storage that are usually poorly utilized as well. They cost more to buy; they take up more data center floor space; they use more energy to power and cool; and they require more staff to manage. All these things are pretty much the opposite of efficiency; yet too often, it was easier to continue to pour money into a suboptimal solution than “bite the bullet” and make things right for the longer term. But in this era of changing consumption models, throwing capacity at everything just won’t work.

That said, savvy storage managers are starting to see some light at the end of the latest recessionary tunnel. They weathered one of the worst economic storms in years, and they did it with flat or reduced budgets, while somehow managing to cope with the digital data explosion.

Now, most IT organizations are looking forward to investing more in storage as they increase efforts to innovate and gain efficiency. ESG research (see Figure 1) indicates that 65% of enterprise organizations and 48% of midmarket organizations expect to increase their 2013 storage spending levels compared with 2012.\(^2\) Overall, ESG

---

\(^1\) This paragraph and a few others are adapted from the ESG white paper, *The Economic and Operational Value of Storage Virtualization: The HDS Perspective*, March 2012.

estimates that the average expected net change in 2013 storage spending is 2.82% for midmarket organizations and 2.93% for enterprises.

*Figure 1. Expected 2013 Storage Budget Levels, by Company Size*

Now the focus should be on ensuring that those budgets are spent wisely—on efficient storage systems. As a reminder, being effective is possible given adequate resources, but being efficient is ultimately more rewarding. Efficiency delivers IT benefits *and* overall organizational efficacy by improving uptime, protecting data better, and speeding new application development just as much as improving costs. We cannot afford to continue bad habits when better alternatives are available.

**It’s Time to Stop Doing Storage the Old Way**

Even though some great “museum pieces” are running in data centers right now, continuing to use outdated storage solutions to address exponential growth is neither an efficient nor a sustainable strategy. Again, the “old way” to meet capacity challenges was to acquire more hardware, thus adding to the data center footprint. It was hardly an elegant solution, but historically, there simply weren’t enough easy-to-use tools available to deal with the storage demands (demands for functionality and flexibility as much as capacity). Although technology improvements make IT tasks easier, they are often outpaced by the growing demands and corresponding complexity of the infrastructure overall.

**A Picture of Inefficiency**

In detail, what does storage *inefficiency* look like? The list is extensive and includes such things as:

- Wasted disk space.
- Untamed, unmanaged growth of cold and archival data on primary systems.
- Limited consolidation of resources.
- Inconsistent I/O response times.
- Long backup and recovery times.
- Trouble meeting recovery time objectives (RTOs).
- Unnecessary energy consumption.
• Increasing RAID rebuild times.
• Inability to deliver high availability.
• Time-consuming storage provisioning.
• Insufficient data protection and security.
• Out-of-control storage costs.
• Duplicate data.
• Inability to prune or remove obsolete data.
• Over- or under-provisioning a storage service.

One of the most visible storage inefficiencies is low disk-space utilization. Utilization rates for most businesses have hovered in the 50% to 60% range for years. Why acquire a 20TB disk array if you’re only going to use 10 to 12TB? And as capacities increase, the pain worsens. Of course, as Table 1 shows, simply deploying a networked storage solution rather than a direct-attached one can improve utilization rates somewhat.3

Table 1. Overall Storage Hardware Utilization Rates Are Higher Among Organizations Citing Networked Storage as Their Primary Disk-based Storage Technology

<table>
<thead>
<tr>
<th>To the best of your knowledge, what is your organization’s overall storage hardware utilization rate?</th>
<th>Organizations citing direct-attached storage (i.e., internal server storage, DAS) as their primary disk-based storage technology</th>
<th>Organizations citing networked storage (i.e., NAS, Fibre Channel/iSCSI SAN, unified storage) as their primary disk-based storage technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>53%</td>
<td>60%</td>
<td></td>
</tr>
</tbody>
</table>


It is in equal measures frightening and entirely understandable that utilizing just over half of their purchased disk capacity has been acceptable to many IT managers for a long time. But today, with the amount of data to be stored growing at an unprecedented pace, inefficient use of floor space and wasting of energy is increasingly expensive. Of course, the percentages shown above represent averages across all respondents, so that 60% statistic is not the maximum utilization that a user might be able to achieve (for example) by deploying a Hitachi Virtual Storage Platform or Hitachi Unified Storage VM solution. The data merely indicates the opportunity that improved efficiency could address.

Imagine if other resources were used as poorly as storage has traditionally been. It’s akin to moving 40% of your just-purchased groceries straight from the car trunk to the garbage can. The money being wasted in unused storage could be working much smarter elsewhere. And the bigger the storage pool, the bigger the inefficiency becomes.

The old ways just won’t work anymore. Unfortunately, years of inefficient practices have produced a sizeable list of storage challenges that are cited in Figure 2.4

---

3 Ibid.
4 Ibid.
**Figure 2. Top Storage Challenges**

In general, what would you say are your organization’s biggest challenges in terms of its storage environment? Which would you characterize as the primary storage challenge for your organization? (Percent of respondents, N=418)

<table>
<thead>
<tr>
<th>Challenge</th>
<th>All storage challenges</th>
<th>Primary storage challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid growth and management of unstructured data</td>
<td>40%</td>
<td>15%</td>
</tr>
<tr>
<td>Data protection (e.g., backup/recovery, etc.)</td>
<td>39%</td>
<td>11%</td>
</tr>
<tr>
<td>Hardware costs</td>
<td>39%</td>
<td>10%</td>
</tr>
<tr>
<td>Running out of physical space</td>
<td>25%</td>
<td>7%</td>
</tr>
<tr>
<td>Need to support growing virtual server environments</td>
<td>25%</td>
<td>5%</td>
</tr>
<tr>
<td>Data migration</td>
<td>25%</td>
<td>4%</td>
</tr>
<tr>
<td>Staff costs</td>
<td>20%</td>
<td>5%</td>
</tr>
<tr>
<td>Management, optimization &amp; automation of data placement</td>
<td>19%</td>
<td>5%</td>
</tr>
<tr>
<td>Lack of skilled staff resources</td>
<td>19%</td>
<td>6%</td>
</tr>
<tr>
<td>Discovery, analysis and reporting of storage usage</td>
<td>17%</td>
<td>5%</td>
</tr>
<tr>
<td>Device management</td>
<td>17%</td>
<td>3%</td>
</tr>
<tr>
<td>File system expansion</td>
<td>17%</td>
<td>3%</td>
</tr>
<tr>
<td>Poor performance (throughput)</td>
<td>15%</td>
<td>4%</td>
</tr>
<tr>
<td>Poor performance (I/Os)</td>
<td>15%</td>
<td>4%</td>
</tr>
<tr>
<td>Power and cooling costs</td>
<td>14%</td>
<td>3%</td>
</tr>
<tr>
<td>Lengthy implementation time</td>
<td>12%</td>
<td>2%</td>
</tr>
<tr>
<td>Poor storage hardware utilization</td>
<td>11%</td>
<td>3%</td>
</tr>
<tr>
<td>Difficulty supporting desktop virtualization environment</td>
<td>9%</td>
<td>1%</td>
</tr>
<tr>
<td>Lengthy storage provisioning time</td>
<td>9%</td>
<td>2%</td>
</tr>
</tbody>
</table>

Meeting the Challenges—What Efficient Storage Looks Like

Fortunately, numerous technologies exist that data storage administrators can implement to improve storage efficiency. The choices range from ultra-high-capacity, secure, low-cost storage at one end, to highly advanced data management functionality and the highest levels of performance at the other. Many of these technologies represent unique levels of storage or “tiers” of the storage hierarchy. Today’s most efficient storage systems integrate SSDs, HDDs, modern tape, automated tiering and migration, load balancing, and deduplication with numerous software and virtualization capabilities to manage the data explosion and drive out complexity. Several of these technologies are “big hitters” delivering superior levels of efficiency. They include:

- Dynamic tiered storage.
- Solid-state disks (SSDs).
- Data reduction technologies—compression, deduplication, and thin provisioning.
- Managing archival and cold data.
- Storage virtualization.
- Shifting of focus from CAPEX to OPEX.
- Storage consolidation (virtualization).
- Energy-efficiency management.
- Optimized transmission of backup data from sources to target(s).
- Effective management, replication, and storage of multiple copies.

Before reviewing the potential value of each of these technical areas from a conceptual viewpoint, it should be made clear that HDS has offerings across the “storage efficiency opportunity spectrum” that cover many of the areas addressed here.

Hitachi Data Systems solutions for primary data storage are Virtual Storage Platform (VSP), Hitachi Unified Storage (HUS), and Hitachi NAS Platform (HNAS). Each is a highly virtualized product that creates a single pool of efficient storage supporting multiple data types (block, file, and content). These systems allow any HDS or third-party array to work as part of one namespace while taking advantage of thin provisioning, advanced replication, logical partitioning, virtual OS connections, and significant scalability and performance.

Scaling to accommodate growth can occur in appropriate tiers while maintaining simple provisioning and central management. Hitachi Tiered Storage Manager provides the data migration function that allows HDS virtual storage to maintain an optimal match of system characteristics with data requirements, enabling easy, nondisruptive storage tiering. Other features such as dynamic provisioning are made available to all storage in the virtual pool.

HDS virtualized, tiered storage simplifies management because the complete, heterogeneous storage infrastructure delivers common storage services. The HDS approach goes beyond consolidating storage to providing an extremely efficient, dynamically allocated set of tiered storage pools. HDS customers are reducing the amount of IT staff needed to manage their storage infrastructures, improving capacity utilization, deferring purchases, eliminating frames, and reducing costs related to hardware/software maintenance, floor space, electricity, and air conditioning. Extending useful life reduces future capital expenditures while improving the performance and availability of data.

Dynamic Tiered Storage and Capacity Efficiency

Tiering has become a storage best practice; the benefits are numerous and significant. Self-optimizing, policy-driven hierarchical storage management (HSM) software automatically moves data to the “best fit” tier available. Optimized tiered storage enables storage managers to place the right data in the right place at the right time. Automated tiering eliminates the need to expend energy on data classification and migration. It improves performance. It increases space utilization, and it reduces the number of storage devices, thus lowering energy costs.

Hitachi Dynamic Tiering is one of many technologies HDS offers to improve capacity efficiency by helping its customers reduce the need for more storage, reclaim capacity, and defer purchases. (Other technologies include thin provisioning, heterogeneous virtualization, intelligent archiving, and write collapsing.)
Judicious Use of Solid-state Storage

Solid-state storage represents the latest addition to the contemporary storage hierarchy. (Although the idea of solid-state has been in play for decades, it was the arrival of less expensive NAND flash and sophisticated tiering and caching software that jointly enabled rapid adoption.)

Solid-state has emerged for high-performance, response-time-sensitive requirements. Those requirements can center on performance as an end in itself or—and this is becoming the most common approach—as a means to an end.

Specifically, the judicious use of a small amount of solid-state can “take the performance strain” as needed and provide for considerable downstream efficiency gains. For example, it can involve the use of fewer high-performance HDD spindles, which then leads to OPEX and CAPEX savings and less of a need for active management.

Most SSDs use flash memory, and to a lesser degree DRAM, to offer the highest IOPS available. Flash memory technology completes operations in microseconds, placing it between hard disk drives (milliseconds) and RAM (nanoseconds) for access time. The appeal of flash-memory-based SSDs is growing: They are nonvolatile, consume little energy, and have a small form factor with relatively favorable pricing levels.

Without the benefit of an SSD option, storage managers looking for speed often short-stroke higher-capacity drives to minimize mechanical disk arm movement (seek time) by allocating only a fraction of the drives’ capacities (typically 5-10%) for performance-sensitive data, leaving the remainder of the drives unallocated.

However, the costly side-effect of short-stroking is significantly reduced capacity: A short-stroked 2TB disk drive might only use 200GB, leaving 1.8GB of capacity unused even though it was paid for and still consumes energy. Judicious use of SSD, on the other hand, supports high-performance application requirements. It improves overall disk—and storage infrastructure—efficiency, yet it constantly maintains, if not improves, storage and application effectiveness.

Data-reduction Technologies Drive Efficiency: Deduplication and Thin Provisioning

Using a data-reduction technology is an excellent way to reduce storage consumed while saving space and money across all tiers. These capabilities address efficiency at the most basic level: They help conserve valuable disk space by removing waste and redundancy. Deduplication, for example, “operates” on the data, trying to reduce redundancies that consume disk space unnecessarily. Thin provisioning enables more efficient disk space utilization by not reserving capacity from the storage allocation pool before there is actual data to be stored. Both techniques can be implemented in such a way as to not affect performance.

Deduplication (and Understanding When It Works Best)

Deduplication can potentially reduce disk storage requirements and expenses more than all other methods. Dedupe’s advantages become especially evident in backup and virtualized environments, where significant amounts of stored data consist of redundant data. The higher the degree of redundancy, the more effective deduplication becomes. A well-executed deduplication implementation can eliminate redundant data and enable customers to reclaim up to 90% of existing storage capacity, extending the life of storage assets by making them more efficient.

The keys to being a well-executed deduplication solution include:

- **Automated operation**—Administrators should not be required to manually schedule deduplication processes or think about when best to turn them on/off.

- **No-impact deduplication**—Deduplication must not affect the performance of the storage system or interfere with the primary file and data storage duties.
• **Maximum efficiency**—Deduplication should not waste a large amount of the capacity it is supposed to be optimizing to track deduplicated data or require usable capacity be set aside as temporary deduplication “workspace.”

**Thin Provisioning**

Thin provisioning is available on most disk arrays. Its purpose is to help IT make sure that overall storage capacity is not “held captive” (i.e., overallocated but unused) by a LUN or volume. Thin provisioning accomplishes this mission by allocating space to a volume only as that space is actually consumed. Some advanced thin provisioning technologies are capable of free space reclamation, which essentially involves returning deleted-space capacity to a global pool of storage. Thin provisioning requires minimal processor and memory overhead, and it is extremely efficient because it does not operate on “in-use” data.

**Managing Archival and Cold Data**

Unstructured, archival, and “cold” (relatively inactive) data is growing faster than any other data classification. CIOs mandating better efficiency often state that “data that isn’t being used shouldn’t consume energy.” Continuing the practice of storing large amounts of cold or infrequently used data on expensive disk is a costly and inefficient strategy compared with the alternatives. That scenario is adding more fuel to the fire of implementing tiered storage—which focuses, as mentioned previously, on getting the right data in the right place to address costs, performance, capacity, and availability requirements.

But to really enact an efficiency strategy for archival and cold data, one must sometimes throw out preconceptions and leverage many other storage technologies, including tiering storage, combining disk plus tape in a strategic “use each for what they are best at” way, and using spin-down disks and data management software.

**Storage Virtualization Increases Management Efficiency**

Most data centers use storage systems and storage network infrastructures from many vendors. Managing each vendor’s system can require having a whole collection of vendor-specific management tools and methods at hand to perform the same tasks.

To address this inefficiency, virtualized tiered storage architectures consolidate heterogeneous storage into a single managed set of tiered pools, providing external capacity with a common set of capabilities—and extending new functionality to older assets. Storage virtualization is an efficiency technique. It can optimize the cost, performance, reliability, and availability characteristics of storage systems from different vendors, matched to application requirements as needed.

Any unnecessary complexity in a storage environment is going to result in higher-than-needed labor costs. Storage virtualization creates fewer opportunities for error. Administrators don’t have to learn multiple ways to execute the same task. Provisioning, tuning, load balancing, troubleshooting, and upgrades happen centrally instead of individually by silo. Storage resources are no longer wasted, and downtime to migrate data is all but eliminated. Virtual storage also reduces maintenance costs. (Software maintenance can often be dropped entirely because arrays inherit the virtualized software features, and consolidation can result in retiring the older arrays.)

**Reducing Operating Expenses: Managing TCO and Shifting Focus from CAPEX to OPEX**

A perception from the old days is that hardware prices ($/GB) should get the most financial scrutiny. It reflects an out-of-date viewpoint that the value of an IT infrastructure lies in its hardware.

The reality is that hardware is the only component of storage-oriented TCO that goes down in price each year, consistently averaging 20% or greater reduction in $/GB annually. (The storage suppliers effectively pass along their economies of scale to their customers.) Conversely, other storage TCO components—software, DR services, energy costs, personnel costs, facilities, insurance, and security—all go up in cost yearly.
Various economic models reveal that over four years, acquisition costs can range from 20% to more than 30% of storage’s total cost of ownership. That means you should multiply your initial capital costs by (up to) five to understand the true cost of storage over its lifetime. Why continue to do it the old way, focusing on the hardware acquisition costs? The total cost of ownership is, appropriately, becoming more important than the sticker price or acquisition cost.

That’s easy to say. Actually doing something about it can be challenging, and it’s the reason for an HDS program called Storage Economics, which applies economic and financial principles to customers’ storage implementations. Since 2001, HDS has been researching, learning, and establishing its thought leadership in this area, and the principles it has developed have been proven time and again. HDS Storage Economics can help by identifying and measuring customer-specific costs, defining economically superior architectures, and aligning them with current and future technologies. Customers end up reducing costs through operational excellence.

Acquisition cost is only one detail in the total cost of ownership of infrastructure. IT and finance managers have heard that fact many times, yet they don’t always know what to do with the information. Figuring out how to locate hard and soft operating costs and drive them out of an IT architecture is complicated—IT organizations are often too busy keeping the trains running on time to think about how to build a better track.

The HDS Storage Economics framework measures and compares TCO among existing and new technologies so that organizations can make much more informed TCO—and ROI—decisions. Working with 34 types of storage costs, HDS can help measure where a company’s costs reside, then design a specific architecture to address them.

**Energy Efficiency Management**

Implementing any of the storage efficiencies above will have a positive effect on energy consumption. Keeping energy costs in check is important, but maybe the biggest reason to manage energy consumption is to avoid maxing out the energy grid serving your data center location.

In some geographic areas, reaching that limit has forced organizations to move to a new data center or add another one. With data center real estate now costing in excess of $500/sq. foot, managing energy efficiency to avoid building a new site takes on real meaning.

Storage managers tend to learn valuable lessons whenever they have struggled to scrape by. Maybe the most important lesson learned is that it’s sometimes better to more efficiently use what’s already installed than continue to add more devices to manage.

Some of the tools mentioned to handle the storage efficiency challenge are already in place. Smart storage managers are now building their own efficiency arsenals to stretch their dollars, tackle their efficiency hit-list, and take as much future complexity out of managing storage as possible.

**Effective Management, Replication, and Storage of Multiple Copies**

Building on the deduplication capabilities within a single storage solution, another aspect of effective copy management includes creating and maintaining multiple copies of data for a variety of business purposes, including test/development, offline reporting, and business continuity/disaster recovery. Those copies, on their own, can represent one of the greatest roadblocks to an efficient storage strategy.

To manage copies efficiently, one should leverage bidirectional block-level replication technologies that can utilize the deduplicated storage pool. By doing that, only the unique data elements, originally created or distilled at any local deduplicated storage, are transmitted to the other appropriate repositories. In this case, efficiency happens in terms of how little space is consumed (regardless of the number of copies) and how little network throughput is

---

5 The fact that HDS has been “beating the efficiency drum” for the past 12 years is noteworthy; many vendors embraced storage efficiency only when the economy tightened. The point is not that this somehow makes HDS “the good guys.” It is more pertinent that the company has been thinking about, dealing with, and perfecting an understanding of storage economics far longer than most other vendors have.
saturated (due to smarter discernment of what should be replicated). But it requires a management tier that understands all of the storage assets across the enterprise.

Table 2 summarizes the foremost “must-have” efficiency features among organizations surveyed by ESG.

<table>
<thead>
<tr>
<th>Must-have storage system features and capabilities, by total disk-based storage system capacity</th>
<th>Less than 100TB (N=190)</th>
<th>100TB to 499TB (N=99)</th>
<th>500TB or more (N=128)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data-reduction technologies</td>
<td>36%</td>
<td>35%</td>
<td>39%</td>
</tr>
<tr>
<td>Storage tiering</td>
<td>30%</td>
<td>30%</td>
<td>38%</td>
</tr>
<tr>
<td>Unified storage</td>
<td>21%</td>
<td>23%</td>
<td>26%</td>
</tr>
</tbody>
</table>

**Additional Observations About the HDS Approach to Storage Efficiency**

HDS defines broad parameters for storage efficiency, encompassing all the applicable tools available across its diverse product lines. Whether customers buy the flagship VSP or the Unified Storage offerings, they benefit from a centrally managed, virtualized pool of storage. Having just one logical pool to manage inherently lends itself to better utilization, but an IT group can also gain efficiencies from sharability. (It’s easier, for instance, to run analytics and migrate across hardware systems.) Also, the IT personnel need only learn one system, which is likely to lead to a virtuous circle of better knowledge and further improvements to efficiency.

We also need to think of “efficiency” in broad terms, not simply according to $/GB. So, for example, the combination of a single pool and smarter IT administrators leads to better, faster, more flexible service levels to the business. Similarly, one might measure efficiency in number of copies of data, reduced network impact by moving less unique data, etc. And those measures and methods directly link to improved competitiveness, customer satisfaction, and other good things.

These same improved business outcomes can be enhanced by improved uptimes (resulting from nondisruptive HDS tools as well as the removal of single points of failure) and the secure multi-tenancy that systems such as VSP and HUS can accommodate. Integration and broad capabilities are the likeliest routes to overall storage efficiency, particularly when cross-enabled through a management lens and software tier that ensures that the storage components and partner-technologies can each add the most value to the overall environment. And in general, “storage efficiency” is a multifaceted and valuable notion. It reflects far more than just the single-feature products that pepper the storage environment overall.

---

The Bigger Truth

Businesses continue to rate managing data growth near the top of their storage priorities for the foreseeable future. Products and solutions now being offered to address that priority are responding to growing demand for capacity aided by storage administrators beginning to understand the heavy burden of using old, inefficient solutions.

Without users and vendors making changes and delivering smarter solutions, the scale and complexity of the problem increases as storage grows. With tightening budgets and external constraints, many IT teams are being forced to do what they should have been doing long before—rethink how they deliver their services. They need to accommodate evolving consumption models in particular, and they need to be more efficient for their companies and their consumers in general. Storage efficiency is a process of continuous improvement, and it may incorporate several projects to achieve the desired goals. The takeaway point is that both the storage industry and end-users must continue to focus more than ever on storage efficiency.

The deluge of data growth has not been tamed even with all the recent technological advances for storage. Vendors must continue to push development of features that allow a storage infrastructure to keep pace with growing data, especially in increasingly highly dynamic IT environments. With armies of mobile devices and virtualized servers that spin up in a matter of seconds, the need to manage primary data, backup data, and the other legitimate copies of data that a business requires has never been greater.

Forget the common old belief that “storage is just storage” and thus is no better than a commodity in terms of being a key part of future IT organizations. If users are to maximize their storage investments and keep pace with business transformation, they must take time to learn about the capabilities of today’s systems and commit to leveraging their efficiencies whenever possible. Otherwise, they will go the way of the dinosaurs.

This is something that vendors also can help to drive through education and training. The first step in addressing any problem is recognizing that you have a problem. The next step is being willing to set aside previous expectations or assumptions in search of a new and better answer than what you thought was achievable before. Otherwise, until the direct linkage between data growth and higher costs can be broken, storage has an unsustainable IT future. Acting within these changing constraints, evolving ones’ previous methods, and imagining new strategies that were unimaginable before will not only enable mere survival by IT leaders, but also should be viewed as a catalyst for IT innovation that will help businesses, teams, and workforces prosper.