

White Paper

Storage Virtualization

What to Know and What to Look For

By Mark Peters

March, 2011

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

Contents

What is Storage Virtualization?	3
Introduction: The “Assumption Gap”	3
Varying Approaches.....	4
Why Would Users Want Storage Virtualization?	4
How We Got Here.....	4
Storage Challenges Users are Facing.....	4
Alternative Approaches	5
Real World Operational and Financial Benefits.....	6
How to Make a Storage Virtualization Decision	7
The Bigger Truth	8

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188.

What is Storage Virtualization?

Introduction: The “Assumption Gap”

All too often in IT generally—and in storage infrastructures specifically—there’s a tendency for an “assumption gap” to exist. This gap is generated by the fact that certain tools and technologies get catapulted to market acceptance—or at least rapid initial traction—so fast that it is *assumed* that everyone not only *knows* what the terms mean, but that everyone also *agrees* what they mean. The problem, then, is not only confusion, but also the fact that many users don’t want to be seen to be the laggards—certainly not from an understanding point of view—and so misunderstandings continue unchecked. As virtualization sweeps through IT, it looks as if a significant assumption gap is opening up around the topic of storage virtualization.

This paper is an attempt to close that gap; in a pragmatic and accessible manner, it covers the real value of storage virtualization (from both business and IT perspectives) and offers explanations of the varying storage virtualization approaches with advice on how to make a technology and vendor decision. Although sponsored by [HDS](#)—which, of course, has its particular attributes outlined—this paper is designed to explain the benefits of storage virtualization in as generic a way as possible and not to merely promote one vendors’ approach.

Virtualization in General

Anyone managing IT—and, specifically, storage—over the last few decades will probably bemoan the complexity that has grown up in the industry; while on the one hand it can be argued that it provides job security, the truth is that it also puts enormous stress on the human resources in IT as much as on the technological ones and it can be a detrimental factor to achieving or optimizing positive business results. Continuing, unabated growth (of storage, applications, and expectations) is only exacerbating the situation and has brought IT to a crossroads. The best answer to address the economic, operational, and financial flexibility required to meet the demands of today’s IT is virtualization, disassociating the system view from the physical assets not just at the margin, but across all facets of IT. Virtualization has the ability (when broadly and well deployed) to ensure that management focus is where it should be: on applications and organizational results, rather than on hardware choices.

If there are any lingering doubts that the data center as we’ve known it is in the midst of a total “virtual” upheaval, let us dispel them: it’s over. The data center of yesterday is gone (or is at least going); the data center of tomorrow is virtualized. Server virtualization, albeit still nascent in its overall capabilities, is not a fad. It is here to stay. It is not simply good marketing. It is logical and compelling, both operationally and financially, and makes business sense so that gradually, and ever-faster, the world of IT is beginning to virtualize its server infrastructure. The more advanced an organization in terms of virtualization deployments, the greater the level of benefit and value that organization can expect.¹

With server virtualization ubiquity a foregone conclusion, the next logical areas to investigate are those which hinder its progress. Storage, at least storage as it has been done for decades, is a significant obstacle that is slowing or stalling the successful and optimal advance of IT. And yet we possess the tools to capably address the issue, with the virtualization of storage being a crucial element.²

In essence, storage virtualization provides *operational fluidity* and *utilization efficiency* by removing the direct linkage between an application, a particular storage type, and even a geographic location on a disk; instead, there is a fundamental storage flexibility that puts the emphasis back on applications, responsiveness, availability, and business value rather than on managing the Ponzi-scheme-like, ever-escalating complexity of traditional storage approaches.

¹ This paragraph—and a few other small sections further in the paper—is an edited and adapted extract from the ESG Market Report “*The Future of Storage in a Virtualized Data Center*,” January 2011.

² It could be interesting for individual users to ask themselves whether in their specific case it is storage itself hindering the advance of IT, or the rapidly increasing growth in data, or even a slow adoption of available technologies. The truth is probably “all of the above” but, although the impact is the same in any case, it may help users to determine the best amelioration.

Varying Approaches

Unsurprisingly, and as the “assumption gap” would suggest, there is no one standard of—or for—storage virtualization. The main approaches, which all have their pros and cons, are:

- **Host-based virtualization models** – where virtualization is essentially a part of the operating system and most storage can have virtualization abilities applied to it
- **Network-based virtualization appliances** – similar to the host approach inasmuch as virtualization is applied to existing storage infrastructure, but in this model, a specific storage virtualization engine sits on the network to orchestrate everything
- **Storage appliances** – where virtualization is built into a specific storage system

The terms “storage virtualization” and “virtualized storage” get applied equally to all of these models. A huge philosophical debate about the semantics is ultimately a pointless exercise because what should be crucial to users is to understand the issues they are trying to address and the positive IT and/or business outcomes they are trying to produce. In other words, understanding the job at hand should be a precursor to deciding on the tool.

Why Would Users Want Storage Virtualization?

How exactly did we get here? Why is storage virtualization so crucial? We know that the implementation of server virtualization invariably requires substantial changes and upgrades to existing storage infrastructure to respond to the stresses and demands that server virtualization places upon them.

How We Got Here

Commercial computing took hold when one single infrastructure stack executed one specific application for one specific purpose. The original mainframe was a glorified calculator. Centralized computing was predictable and controllable, albeit expensive. But it could be managed: one processor system and one IO subsystem.

Decentralized (or distributed) computing was developed largely to try to solve the economic challenges of centralized computing (essentially CAPEX) and yielded low-cost, commodity servers—which we promptly plugged into proprietary, large, expensive, monolithic storage boxes. Servers became cheaper and more interoperable while storage remained proprietary and expensive (though there has generally been a corresponding drop in the price of capacity as processing power has increased). In the old days, the server was the thing that cost all the money. You picked your server by your OS. You picked your OS by your application. Storage was a “peripheral.”

Today, servers are cheap and interoperable while storage is outlandishly expensive, complex, incompatible, and difficult; in many respects, it is the last bastion of IT awkwardness: the peripheral tail wagging the purposeful dog!

Storage Challenges Users are Facing

If it’s not already abundantly clear, it’s worth stating explicitly that storage virtualization is not some technical luxury—on the contrary, it’s all about helping IT to deliver business value and, as such, can be seen as a necessity to address key business challenges. To summarize, the key IT and storage problems faced by many users include:

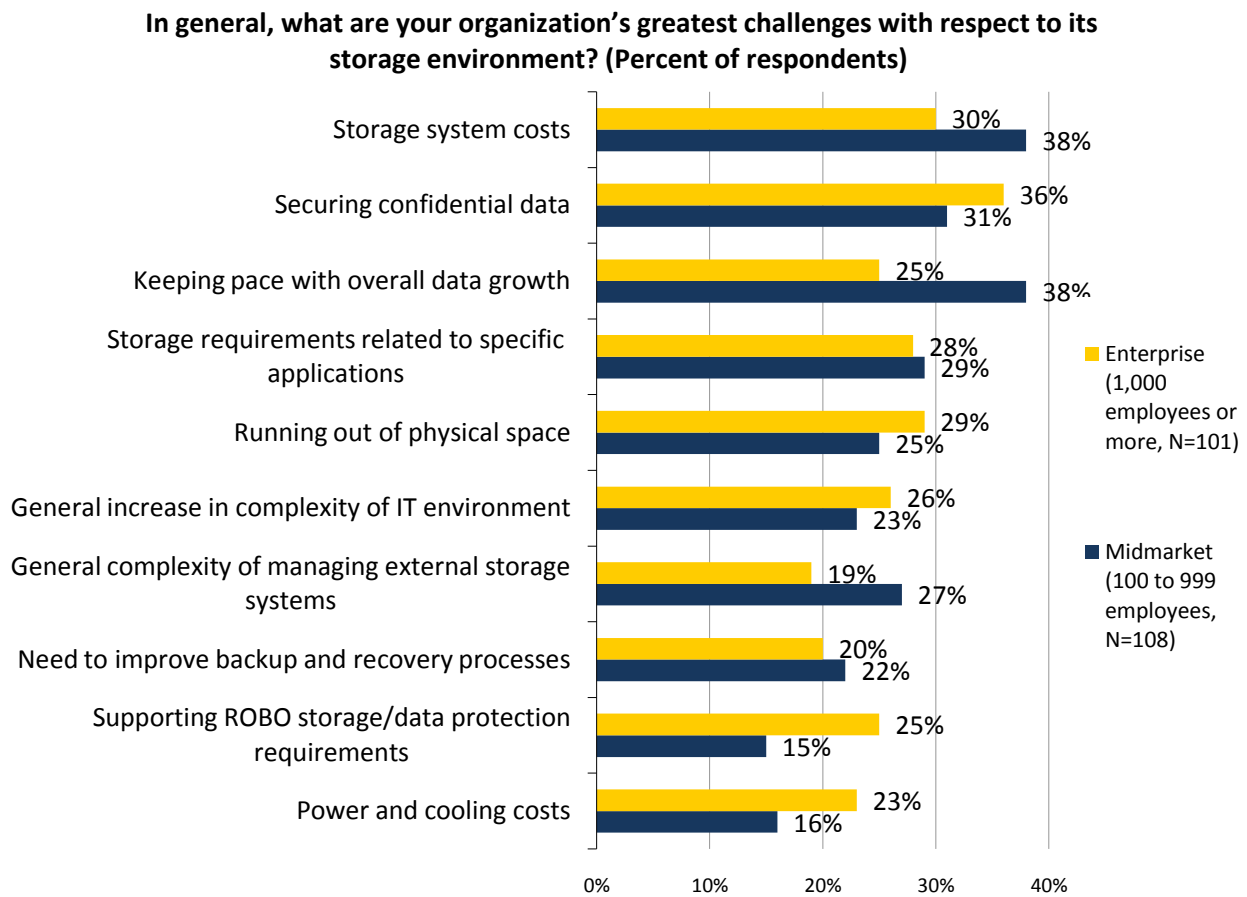
- Capacity growth, together with space inefficiencies
- Virtual machine “sprawl”
- Increasing capital and operational costs, together with constrained budgets
- Inability to scale efficiently

These are not convenient marketing bullets—they are what real users see as their challenges. Figure 1³ shows what IT users say are their greatest challenges with respect to their storage environment. While there are some

³ Source: ESG Research Report, [2011 IT Spending Intentions Survey](#), January 2011.

interesting variations between the mid-range and enterprise accounts, there are also overall similarities: the challenges are growth, security, management, physical space, complexity, and, of course, costs.

Figure 1. Top 10 Challenges in Storage Environments



Source: Enterprise Strategy Group, 2011.

Storage virtualization is a method (a tool, if you prefer) that can simultaneously address—or at least help to address—many of the key *IT challenges* listed. Later, this paper will also show that storage virtualization also addresses the most important business initiatives of cost reduction and business process improvement; indeed, in an IT world consumed by growth and complexity yet constrained by budget limitations, storage virtualization—a marriage of simplicity and suitability—offers one of the few practical solutions.

As already mentioned, storage virtualization and server virtualization are technologies that fit hand-in-glove. Indeed, in recent ESG research focused on server virtualization, storage virtualization was listed as a top-five need to enable more widespread adoption of server virtualization while the number one challenge related to use of server virtualization was the capital cost of new storage!

Alternative Approaches

As with any essentially good general concept in IT, for storage virtualization, there is a range of specific implementations from different vendors. The main choices, with very succinct comments, are:

- Host-/server-based virtualization:** storage control capabilities are provided as a software package that resides on a compute platform and applies virtualization across storage platforms. Interoperability and maturity are the typical areas that prospects want to discuss; users also need to ensure that the availability of compute power is not compromised on a shared host/server.

- **Network-based virtualization appliances:** these essentially take the same approach as the host-based offering, but place the “controllers” (servers) out on the network with the necessary ancillary switches and connectivity. Scaling can be a challenge, but the approach can be excellent in smaller IT organizations.
- **Storage appliances:** this is a variation that, again, combines compute abilities with storage, but with more of the storage control intelligence in special additional switches. It is another workable solution, but obviously it requires more “stuff” and that complexity can lead (albeit not automatically) to islands of data management.

The HDS Approach

Hitachi Data Systems’ approach to storage virtualization is “Array-Based Virtualization” wherein the intelligence resides in the storage controller.⁴ This has the advantage of not requiring additional network components or specialized appliances. Being in the storage system itself also precludes the need for “mapping tables,” and means that storage management stays centralized in one point (especially since the system can manage and virtualize other storage behind it). The latest HDS offering is the VSP, but the storage virtualization has been available on prior HDS models as well—all of which benefit from the vendor’s reputation for rock-solid highly availability and extremely high performance.

Real World Operational and Financial Benefits

While there’s plenty of probability for symbiotic and catalytic value between server and storage virtualization, each can deliver value without the other. All vendors of each type of virtualization can provide plenty of examples and case studies of tremendous IT and business impacts being achieved. Using very colloquial terms, these two tools are as near to a “no brainer” as IT is likely to get; and while the relative value to any given user will vary, the absolute value is logically and invariably going to be greater in bigger organizations. That is indeed what ESG’s research confirms: basically the bigger the IT shop and the more data it has, the more likely it is to be planning to invest in storage virtualization. As Table 1 shows, enterprise users with over 100 TB of data are out-pacing the average for all users in terms of their intent to invest in storage virtualization over the next 18 months by over 42%.⁵ Indeed, for enterprise users, investment in storage virtualization was expected to be second only to investment in backup and recovery tools.

Table 1. Investment Proclivity for Storage Virtualization, by Varying User Profiles

User Sample	Percentage that will Invest in Storage Virtualization in 2011-2012
All users	21%
Enterprise users	25%
Users with 100 TB or more data	30%

Source: Enterprise Strategy Group, 2011.

The benefits that users can gain fall into two main categories:

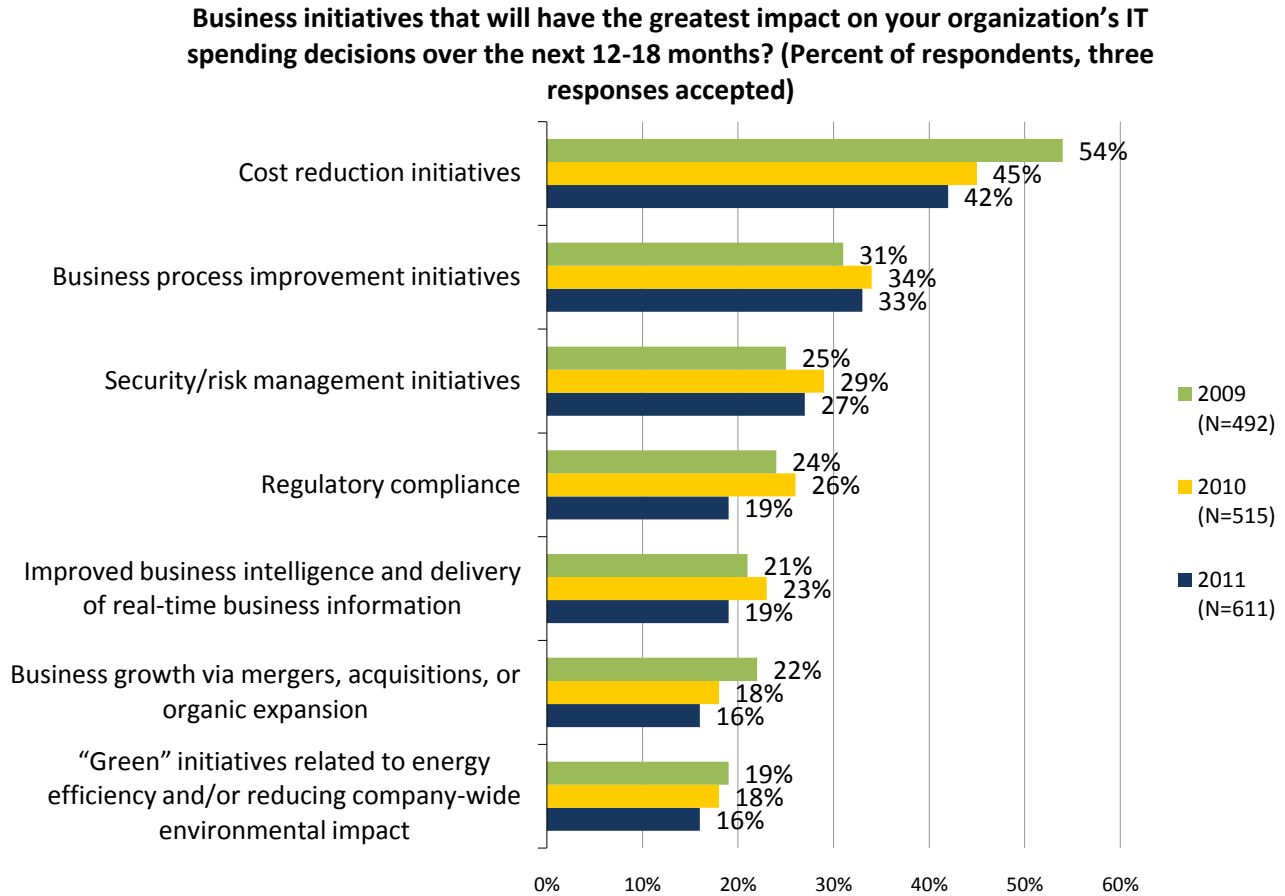
- **Operational** – such as the potential to reclaim space and the likelihood of improving performance to drive better business results.
- **Financial** – such as the possibility to extend asset life, use fewer assets better, and reduce the proliferation of software licenses, reducing both CAPEX and OPEX.

⁴While the purpose of this paper is not to call out or judge specific vendors on all the other virtualization methods mentioned, it is only fair to state that some other vendors—EMC, HP, IBM, and NetApp for instance—also offer versions of this approach (although some of these vendors also offer different storage virtualization options on their various storage platforms).

⁵Source: ESG Research Report, [2011 IT Spending Intentions Survey](#), January 2011.

It should be stressed again that the IT values delivered by storage virtualization are only truly useful if they also apply to improving the user’s business outcomes and ESG’s research shows that the areas that storage virtualization is most likely to affect are exactly those initiatives that businesses are focused on when it comes to deciding their IT investments, as can be seen in Figure 2.

Figure 2. Business Initiatives That Will Impact IT Spending Decisions, Three-Year Trend



Source: Enterprise Strategy Group, 2010.

How to Make a Storage Virtualization Decision

So, now that the general reasons to virtualize storage are clear, how should an enterprise user best approach a buying decision? Like all purchases, it starts with having a clear idea of what is being sought. The following topics are then a reminder and, to some extent, a “caveat emptor question list” which should help any potential buyer to narrow down their options.

1) The basic considerations of how—and what—to buy

- a. What hardware, software, and services are required, not just in terms of “what do you want,” but “what does it take to get it”?
 - i. How feasible—and easy—is expansion beyond the initial system?
- b. What is the method for initial data migration?
 - i. And then for ongoing data mobility?
- c. Take a close look at the overall system management package
 - i. How is the virtual environment managed?
 - ii. Is there storage tiering?

- d. Look at the broader architectural implications
 - i. Networks ... security ... performance?
 - ii. What, if any, interoperability requirements or restrictions exist?
 - iii. Are there open systems and/or mainframe considerations?
- e. What about local and metro data replication?
 - i. Business continuity?

2) Beware of semantics (where the same terms are used)

- a. Many terms that are commonly used have no standard definition, so prudent buyers will dig deeper and compare the details behind the headlines of such things as:
 - i. Ease of use? (which is easy to offer, but far harder to do)
 - ii. Non-disruptive migration? (but to what does this mean non-disruptive: data availability? applications? performance?)
 - iii. Online microcode upgrade? (for example, some systems have no impact, whereas others effectively “borrow” large parts of the cache to achieve the same ends)

3) Compare ends and means (where different terms are used)

- i. Try to get beyond the words to the substance of what is being delivered; users need to investigate and not just “perceive”
- ii. For instance, data deduplication and compression have a similar intent (improving capacity utilization) to zero-page-reclaim, but the latter typically has less performance impact

For all the differences—and they can be substantial—ESG strongly recommends that users do take advantage of storage virtualization. The decision between vendors cannot be prescriptive and will depend on the precise IT situation and needs—perhaps local support capabilities will play a role and even the status of the commercial relationships of a user and vendors may come into play. However, at the enterprise level, storage virtualization offers too many advantages to be overlooked.

The Bigger Truth

A non-virtualized storage environment is almost always going to be a sub-optimal storage environment. After all, these days we can provision a virtual machine (server) for a business in minutes yet we still provision storage in days, weeks, and—often—many months. We can instantly move a virtual machine onto another server and give it more CPU power than entire countries had ten years ago; but if the necessary data isn’t moved with it, it’s a useless exercise. These competing forces cannot continue—operationally or financially. ESG’s position is simple: storage has to cease to be implemented as we’ve known it and must instead become a virtualized complement to the server and network layers of a modern, flexible, effective, *and* efficient IT infrastructure.

The available storage virtualization tools can all provide considerable end-user value; exactly how much value will be dependent on many factors, but potential adopters really need to “get under the covers” and investigate their needs versus the capabilities of different approaches. Hitachi Data Systems commissioned this paper and it certainly has an excellent solution which is thoroughly proven, innovative, and capable—and that absolutely demands consideration—but its motivation in this paper is also somewhat altruistic inasmuch as it wants to promote the overall adoption of storage virtualization; it feels that “a rising tide lifts all boats.” ESG commends that—and recommends that all enterprise IT operations make the investment in some form of storage virtualization since it can deliver the exact types of value that most users are seeking.



Enterprise Strategy Group | **Getting to the bigger truth.**