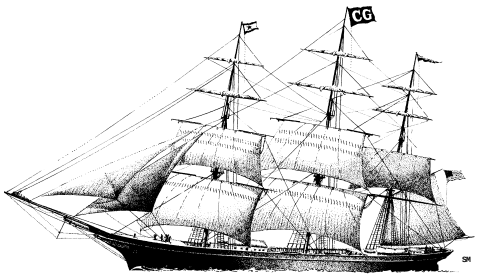
THE CLIPPER GROUP
Explorer™



Navigating Information Technology Horizons℠

Published Since 1993 April 9, 2001

Storage Virtualization in 2001: A Space Odyssey

Analysts: Mike Kahn and Anne MacFarland

Introduction

Storage virtualization is terribly important, possibly the most important information technology concept that you will see in 2001. If you are going to learn one new I.T. concept this year, make it storage virtualization.

With apologies to Sir Arthur Clarke, virtualization is all about managing space. You may think that your storage systems, be they large arrays, network attached storage (NAS), or JBODs[1], are managing data and information, but they are really managing the space on which they are stored. To start, you need to think about the tri-dimensionality of space. Then you need to consider that time (the fourth dimension) changes everything.

Most of us are familiar with the concept of files in general and those in our computers in particular. There are many different kinds of files: documents, spreadsheets, email, databases, etc., but there is a common singularity about them. Each file is unique by name, at least within a specific space, commonly called a folder (or directory). These folders are also unique, at least within their parent space, either another folder (or the "root" of a user's space). Each of us has a predefined space within which we operate. On a personal computer, we may know where that space physically resides (for PCs: the "C" drive), but once you get beyond a notebook, with its detachable and portable nature, you may not really know even where your "C" drive is located. It looks like it is where it traditionally has been, but do you really know? Is it really part of some larger network drive or a block of storage on some really big array, which have been set up to fool you into thinking that it is still on your desk? You are now on the road to understanding virtualization. Most of us think of drives as having names or a "letter" to identify them. Somehow and somewhere, these are "mapped" into physical space on actual hard drives. You've been using basic storage virtualization for years, and didn't even think about it, but your storage administrator has been knee-deep in it for years!

For the moment, think about your storage space as one or more file drawers in traditional file cabinets. Not hard to do, because that is the original model for this electronic representation. Every user has one or more file drawers at their access and disposal. Some are private; some are shared. Some are in the departmental file room, some are in remote offices, and others are in the enterprise repository, located someplace else. In the days of yore, distance was a significant barrier to access. The further away, the longer it took. Today that distance may be still there (the files could be across the country or around the globe), but the time to access them has been reduced to seconds or less, regardless of location.

IN THIS ISSUE

- The Big Benefits of Virtualization..... 2
- Who's in Charge?..... 4
- Whom Do You Trust? 5
- Conclusion..... 6

[1] Just a Bunch of Disks, either internal to a server or externally attached

As a user, you (almost) don't care; you may want it to be faster, but, in reality, it already has been reduced to a "virtual presence" on your desktop. The problems of "making this happen" did not go away; they just have been transferred to someone else: your storage vendor(s) and, most importantly, your storage administrator.

Here's where it gets tricky. In the physical world, we have a limited set of choices for standard office file cabinets – 2, 3, 4 or 5-drawer end-facing or lateral boxes – and we have a few industrial-strength models, like the mechanically-moving ones that store thousands of files, as you might find in a hospital. In the electronic storage world, there are a plethora of types and models and manufacturers, from the 4-drawer-equivalent to storage farms. Each storage system, regardless of configuration and capacity, thinks that it is in the real estate business, providing cubic space for your files.

So when you look across your enterprise storage real estate, you may find buildings of different architecture, capacity and vintage, in many locations, all of which somebody needs to manage. Your IT department is your real estate department and it employs space managers, called storage administrators, to get the most out of this physical asset, in terms of space optimization, tenant satisfaction, and cost of service delivery.[2] Storage administrators have many things to do, but, today, they are typically driven by the challenge of where to put it all.

For simplicity, think of all of the storage as residing on one large plot of real estate, populated by many different kinds of warehouses and structures to hold data. Your storage administrators not only have to worry about where to place your data, but how to give you a pathway to it, and who is responsible for directing you to it. Think roads, street addresses, loading docks, elevators, room numbers and file cabinets. Then they might have to worry about moving it from your warehouse building to another building, maybe of a different type, because of priority-need-to-access, renovation, security, high availability, and the like. And, because you do not like putting all of your eggs in one basket, maybe you want to have copies of some of these files in another warehouse on the block, or much farther away. Your storage administrators have to make all of this happen, and they have been doing this, at least for mission-critical data at larger enterprises. However, this is an increasingly burdensome load that they are carrying. Too much to store and not enough administrative capacity to keep doing it the same way. There is no choice; your enterprise has got to change its warehousing scheme.

Right now your enterprise has a number of storage administrators, each with their own space to manage and their own expertise in making it work in the building(s) that they manage. But remember that all buildings are not of the same architecture and that their colleagues may have different kinds of warehouses to manage. So you have a bunch of physically-oriented (space) storage specialists (by manufacturer, connectivity, architecture, and/or tools-of-the-trade), doing your virtualization for you. Wouldn't it be better for the enterprise to make this simpler? You probably won't see the difference, except possibly in lower costs, but wouldn't it be easier to manage if all of the real estate could be managed from a single perspective?

With storage attached to a network (NAS or SAN), it is possible to think of storage not by the box or disk, but as a plot of real estate or a piece of three-dimensional space. You might quantify it in megabytes or gigabytes and reference it as a disk or a logical unit called a LUN. With traditional methods, you may actually know where something is located in space. With virtualization, your storage blocks can be moved around. You don't know and don't care where they are located or that they

[2] You may be outsourcing part or all of your storage real estate and/or management to a storage service provider (SSP), but the model still works. It's just that the people doing the virtualization may be outsiders. More on this in a future bulletin.

Copyright © 2001 by The Clipper Group, Inc. Reproduction prohibited without advance written permission. All rights reserved.

may have been moved, except for characteristics of quality of service that you have requested, either directly or indirectly.

Eureka! That's storage virtualization circa 2001. In a nutshell, storage virtualization is the transparent use, storage and management of enterprise data by aggregating it into a collective pool and doling it out as "virtual space" as it is needed. Management is the key, because, as we explained earlier, enterprises have been virtualizing storage – without the knowledge of users - for many years.

The Big Benefits of Virtualization

Now that you understand the concept of virtualization of storage, it's time to think about the benefits to an enterprise, which we group into three categories: economic, greater agility, and administrative.

Economic Benefits

While you will always need to have headroom for growth, you do not need it on every volume of storage in your possession. If you reclaim space that is sitting unused on your existing arrays and are able to put most of it into production[3], then there is an immediate return on investment equivalent to the cost of storage that you didn't have to buy. We are not talking about small change here, since estimates of over-provisioning can run well over 100%. Storage pooling puts more of your storage to work for you, and gives a substantial and immediate return on investment.

Clearly it is less wasteful, but the nature of storage virtualization is also more open-ended than fixed-assignment schemes, making it much easier to add capacity to the pool and have it automatically put to use where it is need. The difficulties of adding storage or changing drives on the fly are major reasons why administrators over-provision in the first place.

Not only is it open-ended but, by its very nature, it tends to be self-optimizing, moving blocks and/or files behind the scene to achieve policy-based goals such as response time to delivery and prioritization of access. This gives the administrator a proactive handle on thorny quality-of-service issues.

Many of the today's critical storage management functions, such as making copies on the fly or remote copying, are now separately priced software packages, sold by the storage vendors on a storage subsystem basis, and run inside the storage subsystem. Remember that the more storage subsystems that you have, the more you have to pay for this software. Because of the centralization that comes with virtualization, this same storage management functionality may prove to be much less expensive.

Some of these functions are truly done in hardware today. Most of storage subsystems have large amounts of cache or temporary memory that speeds up the flow of data into and out of the subsystem. Under many virtualization schemes, this is done in a near-commodity storage server or appliance, where the memory costs less and it can be shared more easily among different files. Some go further and do RAID management within the storage server as well. Storage virtualization should allow you to consider less-costly storage with fewer high-cost features, for a significant net savings.

So what are you going to do with all of the money that you are going to save? Of course, you could buy more storage, and you will. But you should also consider doing more than you

[3] All volumes of storage have unallocated space and/or space reserved for growth of data and files. This is normal and necessary. The excessive cost comes from each volume having that "slop at the top", which adds up to a lot of excessive capacity held for future use. Reasons for slop at the top include the difficulty of adding storage capacity to an array and assigning more storage to a server or application. With storage pooling, the reserves are consolidated and, usually, significantly much less "slop" is required for the total storage pool. This can result in significant use of previously unused storage capacity, not to mention eliminating the need to buy more slop capacity. The other reason is that storage comes in prescribed units, like 2, 3, 4 and 5-drawer file cabinets, which may be larger than you currently need. Thus you may be able to procure more granularly and buy the kind of storage that meets your (policy-based) needs.

Copyright © 2001 by The Clipper Group, Inc. Reproduction prohibited without advance written permission. All rights reserved.

deem affordable and reasonable today. Almost every organization has information that is underprotected, because the cost of protection didn't seem reasonable. With virtualization, you can protect more (more frequent backups or more remote copying) or do more (more copies for sideband use, say, for a data mart).

The Benefits of More Agility

If enterprises stood still, even for a moderate period of time, storage administration would move from difficult to do-able. But the pace of enterprises is increasing, more data is being processed every day, new transactions are regularly being added to the mix, compounding the challenge. Your competition does more, your customer wants more, and you must deliver more. You must be more responsive exactly when change is accelerating, data and transactions are increasing, and workloads are becoming more unpredictable. Storage virtualization significantly increases an enterprise's agility to do what needs to be done. No longer are you tied to the physical world of your storage. As needs change, say for a new server, you can just allocate the old storage to the new server, without rewiring and reinitializing, or you could move the files from one storage subsystem to another behind the scenes, without the servers or users being aware.

Quite apart from IT considerations of using media more efficiently and allowing data access more gracefully, virtualization lowers the barrier to organizational change thrown up by the complication of repurposing and reassigning storage assets. With storage assignable on the fly, files and data can be put wherever needed, as storage needs change. The Net, wireless, and other advances in communications, not to mention mergers and scale-backs, have pushed a litany of changes in business, from the addition of new channels and new forms of product support to re-qualification and rededication of work forces, e-procurement, remote workers, and other virtual relationships. Temporary alliances have become essential to getting products to market quickly. Management has high expectations on how quickly an organization can and will evolve, sometimes with inflated ideas of what is possible. All of these demand an agile enterprise: storage virtualization adds to that agility.

Administrative Benefits

With all of this said, the biggest payback of storage virtualization may be on the people front. Fewer administrators can manage significantly more storage under a virtualization model. There is less to do because more is done automatically. There is now a single point of administration that is mostly-to-completely independent from the storage subsystems. Everything is easier with virtualization, when all of the storage looks the same, due to pooling. With policy-based administration, it even gets easier.

Who's In Charge?

Your enterprise is being bombarded with vendors, some traditional and some new, hawking storage virtualization. It's their latest thing since "sliced bread". Here's how to process some of what they are saying.

- Each of the vendor's wants to be in charge of your virtualization effort. This puts them at the center of the your storage strategy.
- Suppliers want to establish themselves as the standard for virtualization, although none exists at this time. Each says that they are "open", more or less, but in reality they are open as long as you hire them to be your warehouse manager and, in some cases, buy the warehouse buildings from them. Today, openness has more to do with connectivity to servers (operating environments), network connectivity (SCSI versus Fibre Channel versus IP), and storage systems.
- Each virtualization scheme has a locus of intelligence. The reality is

Copyright © 2001 by The Clipper Group, Inc. Reproduction prohibited without advance written permission. All rights reserved.

that, even though servers become much less responsible for managing the file systems under virtualization, the intelligence has just been moved to one or more specialized virtualization servers. Think of these as warehouse directories, where the user's file name is mapped into a building, floor, room, file cabinet and drawer. Realize, however, that the value of virtualization is that the user doesn't need to know and doesn't really care where his file is stored, as long as he can reference it the way he wants.

The proverbial block of warehouse real estate described above might be run by a storage manager, whose job it is to know where everything is stored. Remember, the storage manager's responsibility is to manage the storage space, not the contents therein. He might use a centralized directory for all of the warehouses, or there might be several zone directories on the block (each covering a number of buildings or there might be one in each building. The central directory and zone directories might called "out-of-band", because the directory sits in some third location, and you can't get to the right building without going to the third location first. On the other hand, having a directory in each building is "in band", but you have to know which building has your stuff. The same terminology applies to virtualization. Some virtualization directories are kept on in-band devices and others are externally located out-of-band; each with different operational characteristics and life-cycle costs.

For some schemes, the work may be done on traditional servers; for others, it may be done on a dedicated or special-purpose device. The mapping has to be done somewhere, and you are going to pay for those work cycles. Deciding whether you want your virtualization done in a general- or special-purpose server, a network device (such as a router or switch, or even "in the network fabric"), or in a storage system, is serious architectural decision. There is no one right answer; it depends on a lot of enterprise and I.T. infrastructure variables. You may need to seek outside, independent help on this one.

Under any virtualization scheme, you need to ask some tough questions.

- Will you be able to use the current storage buildings that you now have, those warehouses called storage systems?
- Will you support different vendor's storage? (Will one forklift work with all of your pallets?)
- Will the benefits of virtualization apply only to new warehouses (storage systems) procured for the purpose of virtualization?
- Will you need new roads (network infrastructure) or does the vendor's offering support what you already have?

There are many trade-offs to be considered with virtualization. Regardless, you are going to have to provide (and pay for) this entire virtualization infrastructure, regardless of the scheme. The good news is that it should not take long to recover this investment.

Whom Do You Trust?

This is not an easy question to answer. Access to data is core to most businesses, and storage virtualization is one project whose implementation has to go well. By addressing the questions outlined above in ***Who's in Charge***, you can determine your hot buttons and select vendors that address your concerns. This will winnow out some. You then have to address your own characteristics. It all boils down to "whom do you trust" and why.

The risk averse may prefer to buy from established vendors, even though they know that they will pay more and potentially have fewer alternatives. The value buyers may look to the new players for more functionality and

Copyright © 2001 by The Clipper Group, Inc. Reproduction prohibited without advance written permission. All rights reserved.

connectivity, and maybe at a lower cost. If you need help in implementing virtualization or at least getting started, you should check out the service providers and their cost models. In the short run, it's making the comfortable choice. In the long run, it's about the total cost of ownership.

One of your early steps should be internal education. Most IT projects that fail do so because of a lack of internal buy-in. It is important that storage virtualization not fall prey to being viewed as another techie diversion. Management needs to understand the capabilities empowered by virtualization, including an enterprise's ability to respond to changes, and the tangible economic benefits, including leveraging downward the cost of storage and administration.

Conclusion

It's all about managing your enterprise's most valuable real estate, the information space. Virtualization is a strategic opportunity for your enterprise to take control of your enterprise data space without having to sweat all of the day-to-day details, by moving to a higher plane of understanding and existence. Done right, your virtualization space odyssey should deliver a handsome return on investment, in terms of the impact to the enterprise and the maximization of your information technology budget.

Storage virtualization may be your biggest opportunity to achieve ***teraproductivity*SM, or making big things very productive for your enterprise.** That's your challenge — go for it!

Copyright © 2001 by The Clipper Group, Inc. Reproduction prohibited without advance written permission. All rights reserved.

About The Clipper Group, Inc.

THE CLIPPER GROUP, INC. is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. Our capabilities are augmented by a team of staff consultants with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at (781) 235-0085 and found on the Internet at www.clipper.com.***

About the Authors

***Mike Kahn* is Chairman and a cofounder of The Clipper Group.** Mr. Kahn is a thirty-year veteran of the computer industry. For the vendor community, Mr. Kahn specializes on strategic marketing issues, especially for new and costly technologies and services, competitive analysis, and sales support. For the end-user community, he focuses on mission-critical information management decisions. Prior positions held by Mr. Kahn include: at International Data Corporation — Director of the Competitive Resource Center, Director of Consulting for the Software Research Group, and Director of the Systems Integration Program; President of Power Factor Corporation, a Boston-based electronics firm; at Honeywell Bull — Director of International Marketing and Support; at Honeywell Information Systems — Director of Marketing and Director of Strategy, Technology and Research; with Arthur D. Little, Inc. — a consultant specializing in database management systems and information resource management; and, for Intel Corporation, Mr. Kahn served in a variety of field and home office marketing management positions. Earlier, he founded and managed PRISM Associates of Ann Arbor, Michigan, a systems consulting firm specializing in data management products and applications. Mr. Kahn also managed a relational DBMS development group at The University of Michigan where he earned B.S.E. and M.S.E. degrees in industrial engineering.

- ***Reach Mike Kahn via e-mail at MikeKahn@clipper.com* or via phone at *(781) 235-0085 Ext. 21.***

***Anne MacFarland* is Director of Enterprise Systems Research at The Clipper Group.**
Ms. MacFarland specializes in the strategic solutions being offered by enterprise systems and storage vendors. She joined The Clipper Group after a long career in library systems, business archives, and research, including work for Connecticut Historical Society, Stowe Center, Aetna Life and Casualty, and Travelers Insurance. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

- ***Reach Anne MacFarland via e-mail at AnneM@clipper.com or via phone at (781) 235-0085 Ext. 28.***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, and **The Clipper Group Observer** are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and "teraproductivity" are service marks of The Clipper Group, Inc. The Clipper Group, Inc. reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc. has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.

Copyright © 2001 by The Clipper Group, Inc. Reproduction prohibited without advance written permission. All rights reserved.