

White Paper

NetApp Delivers Efficiency at Scale

By Mark Peters

January 2013

This ESG White Paper was commissioned by NetApp and is distributed under license from ESG.

Contents

The Challenge of ‘Big’	3
The Market Situation.....	3
NetApp Users Are Beginning To Buck the Trend.....	4
1) Market Growth Accelerates	5
2) Is Efficiency At Scale Trumping Human Skills?	7
3) Data Protection at Scale	7
4) Big Data Further Drives the Need for Efficiency at Scale	7
NetApp Offers Rich <i>Efficiency at Scale</i> Management.....	8
Scale - (Way) Out with Clustered ONTAP	8
NetApp’s Advantage: Robust Storage ‘Efficiency at Scale’ Technologies	9
The Real World: A Case Study, and More.....	11
PeakColo	11
More to Do: Efficiency at Scale Is a Journey, Not a Destination	12
A Final Reminder....and a Recommendation	12
The Bigger Truth	14

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.

The Challenge of 'Big'

IT is going through an evolutionary transformation fueled by technology, new business factors, and unprecedented growth. Some of the major trends contributing to this transformation include IT service automation, converged infrastructures, virtualization, private and public clouds, the rapid adoption of many types of flash storage, both scale-out server and storage architectures, and a common management template for different technology systems. While new technologies in place are enabling organizations to transform their data centers without compromising performance, scalability, or availability, the combined issue of size and increasing complexity is still paramount for many IT organizations. Whether an organization is already a large enterprise or midsized and growing, the issues are either “in your face” or imminent. Smaller “shops” need to be planning and deploying scale-optimized infrastructure as everyone will most likely wind up at the petabyte level eventually. Indeed, as we shall see later in this paper, there is a certain “tyranny of big numbers” to deal with—meaning that many things (such as efficiency and manageability) simply get harder at scale. In fact, these things can get harder at a rate that makes fixing them almost harder still! There is irony in this: IT users don’t reach out to become more *efficient* because they perceive (or at least believe) that it’s *inefficient* to get there! Or perhaps there’s a concern around “where do I start?” It’s interesting that growth, backup, and efficient management have always been the basic “blocking and tackling” issues in storage, yet they’re still the ones that are most challenging.

With all sorts of demands on storage increasing, effectively managing the way that ever-increasing amounts of data are created, stored, and accessed only intensifies the challenge of *managing at scale*. To address data growth and yet not interrupt business operations is a huge part of the equation. The ability to rapidly deploy storage and IT resources to meet increasing demand is mainly a function of scalability. Size and scale matter a lot and larger organizations need to have even greater capabilities (while midsized organizations need to be prepared). But the question remains: How should organizations optimally impose efficiency on a long-term challenge that is at once stubborn *and* dynamic? While it is easy to agree that much of the traditional approach is unsustainable and that “something must be done,” the issue is no longer akin to the old adage of changing the oil during a car race...once everything is at scale, it is more like changing the transmission in all the cars, *and* resurfacing the track at the same time!

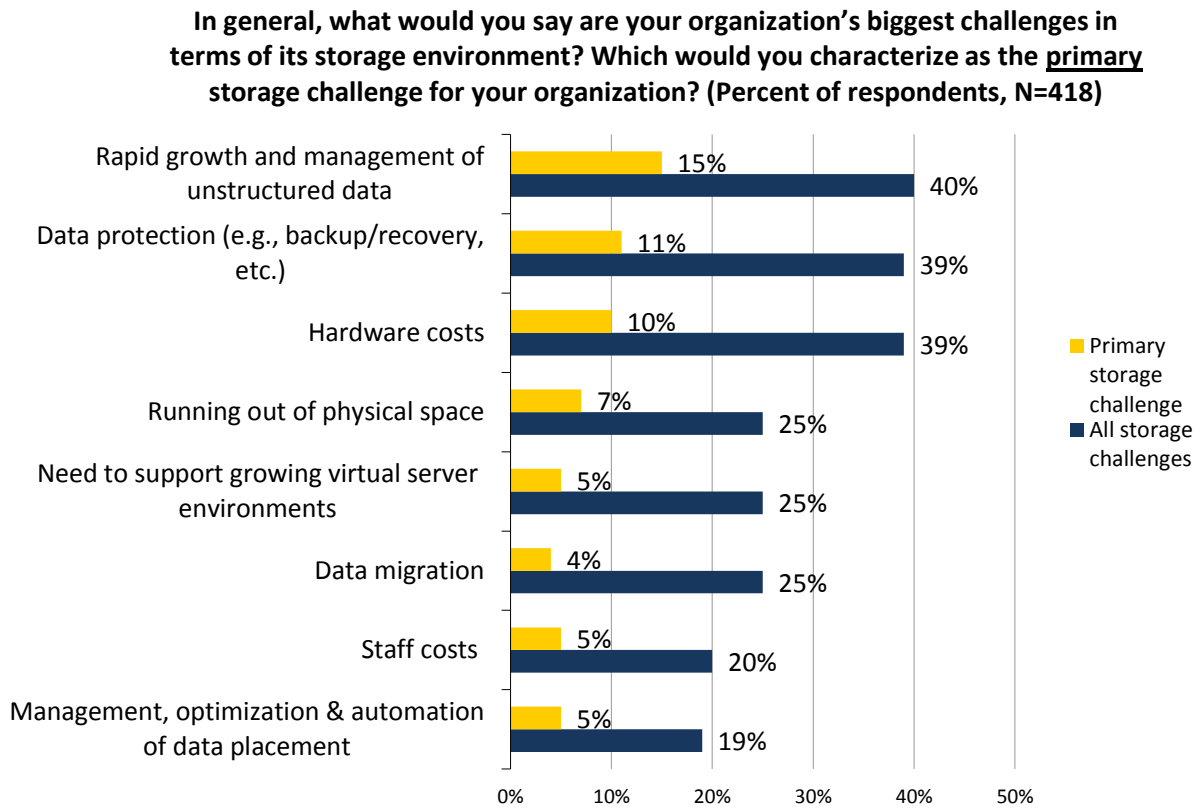
The primary objective of *efficiency at scale* is to create a dynamic, non-disruptive environment that supports data growth in such a way that system capabilities remain as balanced as possible while, in parallel, meeting constantly changing and more demanding business requirements. A huge shift in IT is underway from an approach of dedicated architectures to more flexible, efficient, and shared architectures. [NetApp](#) has been working on the challenges of scale—and of providing efficiency at scale—for years, and it has developed some strong and commendable abilities to deal with “big.” This paper examines the issues, and NetApp’s answers.

The Market Situation

As computer and storage prices drop, and performance and capacities continue to increase, low-cost “commodity” systems are increasingly being used for high-performance and high-capacity workloads. In the past, this could only be handled by scaling up to much faster storage devices, or high-performance computers. With scale-out abilities, many small, low-cost elements can be configured to create an aggregate storage or processing pool. The scale-out model has created an increased demand for shared/clustered/grid data storage, storage virtualization, and improved data protection services. Another consideration for managing data growth and storage technology is the necessity for increased skilled human interaction. Today, IT administrators (aided by the advanced and often automated functions in modern storage systems) can manage far more data and storage per person than was considered possible just a few decades ago. While this has helped to somewhat mitigate the number of personnel that are needed, the fact remains that it is hard to avoid increasing headcount as a method for managing nonstop data growth. In addition, the linkage between the rate of storage-price decline and the rate of capacity-demand growth has been broken. Storage demand is now growing faster than the price is declining, so—assuming limited budget growth—the current storage model is unsustainable without substantial changes in the ability to efficiently manage storage at scale. Whether via anecdote or as reported by respondents to an ESG research survey (see Figure 1), we know that data growth remains the prime storage challenge; as

shown, data protection and the associated hardware costs are also key challenges.¹

Figure 1. Major Storage Challenges



Source: Enterprise Strategy Group, 2013.

As a pragmatic tool that tries to address this challenge of growth, scale-out storage has not only made its way beyond markets typically requiring bandwidth and performance scaling, and into conventional IT environments, but is also demonstrating extremely fast rates of compounded annual growth. Looking at the practical impact of this, such growth can represent both a massive operational headache and a massive budgetary need. While some of the budget pain might be addressed by natural budget increases and/or natural per-GB price declines, there is clearly still a need for dramatically improved efficiency to make such growth and scale practical.

NetApp Users Are Beginning To Buck the Trend

Efficiency is an over-used word in the world of IT: It’s become like quality, better, reliable, faster, cheaper, etc. However, *efficiency at scale matters*. Whether you are a large enterprise or a growing mid-sized organization, you could be looking at “recovering” (or not having to initially buy) significant percentages of your storage: NetApp analysis of its base has actually found that customers with more data are achieving 16.4% more savings than smaller environments (those under 50TB). Such details are known from aggregate analysis of the impressive My AutoSupport tool²—a web-based application that provides customers with insightful information designed to analyze, model, and optimize their storage infrastructures.

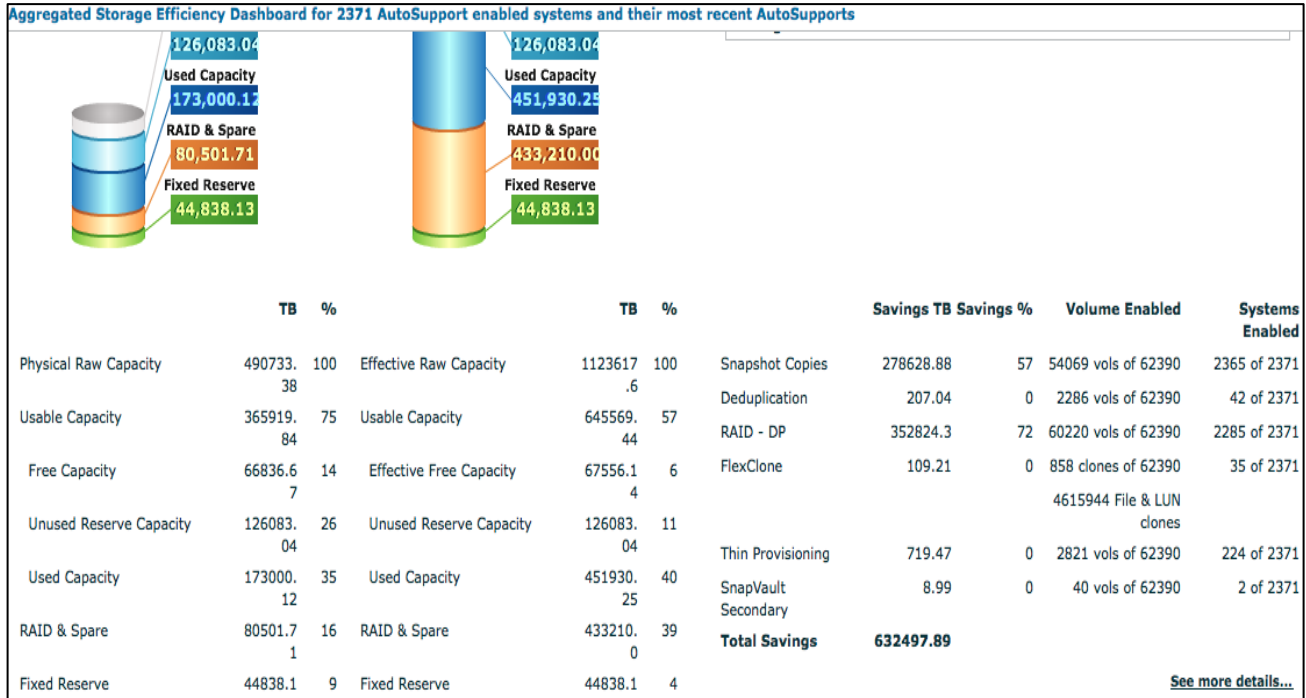
My AutoSupport evaluates system metrics and has a storage efficiency section which shows the capacity used and saved that can be attributed to each of the NetApp efficiency features installed. My AutoSupport indicates the percentage of storage saved by NetApp’s Snapshot Copy, Dedupe, Thin Provisioning, Virtual Cloning,

¹ Source: ESG Research Report, [2012 Storage Market Survey](#), November 2012. (All other ESG references and charts in this White Paper have been taken from this research report unless otherwise noted.)

² For more information go to: <http://www.netapp.com/us/support/autosupport.aspx>

SnapVault, and RAID-DP features. This is a level of storage insight not readily available from other vendors’ measurement systems. Figure 2 shows a sample of the dashboard. Overall, the biggest savings in this example can be seen to come from Snapshot (57%) and RAID-DP (72%), and individual users can readily see how many TBs are saved by all of the efficiency features they have installed (of course, since NetApp has some 170,000 customers there’s a lot a variation, which is why the AutoSupport tool can be so valuable to individual users).

Figure 2. NetApp “My AutoSupport”: Where—And How Much—Am I Saving?



Source: NetApp, 2013.

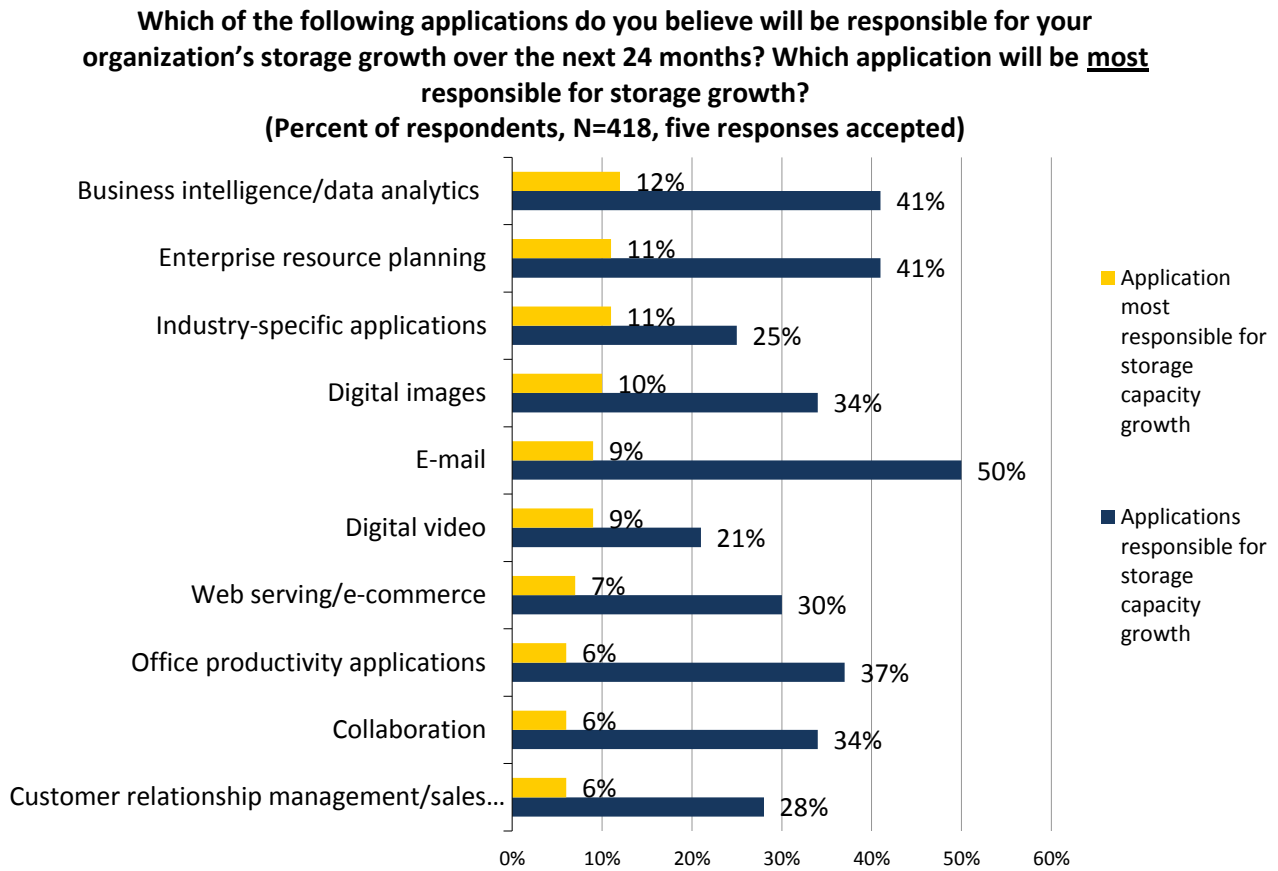
They key point here is that NetApp is actually showing its customers how to buy less storage! By using the AutoSupport tool, users can see precisely how to improve their storage efficiency (i.e., which functions are being well used, which are not). To use a “human efficiency” analogy, there is a big difference between buying a treadmill, actually *using* a treadmill, and using a treadmill optimally.

Understanding these results is crucial because four major issues are conspiring to make the drive to efficiency at scale increasingly more challenging and, thus, more important. The issues are as follows:

1) Market Growth Accelerates

What’s driving all the data growth? ESG’s research in Figure 3 highlights the applications that are the primary drivers of storage growth pressing the obvious need for more storage efficiency at scale *without* compromising effectiveness. The pressure of big numbers—the doubling (or more) of capacity—simply means more of “everything” at increasing scale. Some users know they’re going to eventually be consumed by it all, but this gets ignored because the issue is generally hard to fix and many don’t know how to begin their attack. As mentioned previously, the irony that IT groups don’t reach out to become more efficient because it’s inefficient to get there is ever present—but it is also an unsustainable strategy. The challenge facing most enterprises today is that data growth—and file growth in particular—is, effectively, out of control. This is a major factor driving scale-out NAS and its granular, “plug-and-play” growth capability to break out of specific vertical use cases and become a major segment of the overall file-based storage solution market.

Figure 3. Applications Responsible for Storage Growth



Source: Enterprise Strategy Group, 2013.

As part of the sheer challenges of scale, it is also interesting to analyze applications by company size. As shown in Table 1, larger enterprises are more likely than their midmarket counterparts to see data growth associated with business intelligence and data analytics (48% enterprise, 28% midmarket). These are, of course, applications that tend to *drive random storage performance and flexibility harder* than the more straightforward and linear capacity/performance growth associated with the more “regular” applications on the list. It is another reminder—if one were needed!—that efficiency at scale cannot come at the cost of any diminished effectiveness.

Table 1. Applications Responsible for Storage Growth, by Company Size

Which of the following applications do you believe will be responsible for your organization’s storage growth over the next 24 months?		
	Midmarket (100 to 999 employees, N=141)	Enterprise (1,000 or more employees, N=277)
Business intelligence/data analytics	28%	48%
E-mail	58%	46%
Web serving/e-commerce	23%	34%
Office productivity applications	42%	34%
Video surveillance	20%	14%

Source: Enterprise Strategy Group, 2013.

2) Is Efficiency At Scale Trumping Human Skills?

Of course, data growth invariably results in a need for more overall storage capacity and raises another consideration for managing such data growth: the necessity for additional skilled human management interaction with storage systems. The link between overall storage capacity managed and the number of people required for the task still exists. Specifically, ESG research³ indicates that organizations with at least 500 TB of total storage capacity are twice as likely as those with less than 50 TB (42% vs. 21%) to hire additional storage administrators in 2012 (see Table 2).

Table 2. Organizations with More Storage Capacity Under Management Are More Likely to Hire Storage Administrators in 2012

	Total storage capacity		
	Less than 50 TB (N=42)	50 TB to 499 TB (N=43)	500 TB or more (N=33)
Percent of organizations planning to hire new storage administration staff in 2012	21%	33%	42%

Source: Enterprise Strategy Group, 2013.

This reinforces the premise that data growth is still outpacing the abilities of much of the “regular” technology deployed to manage it. The amount of storage an individual can manage these days far exceeds what once was the case, but the fact remains that it is still hard to avoid increasing headcount as a method for managing data growth without a management tool like My AutoSupport.

3) Data Protection at Scale

As the demand for storage rises, demand for ways to protect the stored data increases as well. Users are often unsure about how to provide data protection at scale. Having a goal such as “mission-critical applications are always protected” can be more hope than substance. *Remember: Hope is not a strategy!* Traditionally, when data protection is mentioned, most people immediately think of tape for backup. However, over the past several years, some emerging technologies such as replication, synchronization, and snapshots have established their place within the data protection realm. While they offer a potentially more expensive solution, these emerging technologies also yield a much faster RTO, which is of course a requirement for most mission-critical applications. Snapshots offer quick user-level restores without the need for administrative assistance, while synchronization and replication provide business continuance and offsite disaster recovery with fast RTOs.

The upper tier of organizational management doesn’t always realize that their mission-critical, revenue-generating, customer-facing applications may be insufficient to run the business by themselves. Furthermore, there is a new and growing concern to protect “compliance data” (which can easily make up over 20% of all data in a business). Many businesses choose to ignore adequately protecting this data, thus adding to an expanding list of unmet data protection requirements. In other words, the need is not only for efficiency at scale, but also to provide more efficient data protection at scale.

4) Big Data Further Drives the Need for Efficiency at Scale

The term “big data” typically refers to significantly large sets of data as well as to the tools organizations must use to create, manipulate, and manage them. Data sets become big data when their size and breadth grow such that it becomes hard to derive business insight from them without specialist tools. Much of this growth is in the form of file data. When a company is amassing data assets with rapid growth rates, the sheer file size or volume of file

³ Source: ESG Research Brief, [Does Data Growth Still Outpace The Ability Of Advanced Storage Functionalities To Tame it?](#), July 2012.

storage and access operations creates new performance challenges that traditional architectures often cannot scale to meet.

Conventional NAS was a perfect way to store and protect many small files. If an abundance of additional small files were being created, then a scale-up expansion was a fine way to support the growth. But today, big data is proliferating. Even smaller enterprises are now moving large data sets around frequently. The biggest of the big data users commonly create and move individual files that are each many terabytes in size. Simply adding resources to storage, bandwidth, or computing power no longer addresses all of their needs. In the era of big data, users need their NAS environments to scale not just in capacity, bandwidth, and compute strength, but also in throughput. The sheer volume of storage and access operations underway is creating performance bottlenecks that traditional architectures can't always meet. Big data files only add to the stresses.

NetApp's high-performance scale-out systems, with their optimized controllers and advanced clustering capabilities, may be the answer for companies that are rapidly generating extremely large files. Let's look at the latest NetApp solution suite to see how it can help to successfully meet these challenges by delivering a broad and deep range of capabilities—including scalability, management-ease, data protection, and big data—that can be summarized as “efficiency at scale.”

NetApp Offers Rich *Efficiency at Scale* Management

An optimized, next-generation storage system will need to scale out *and* up; support multiple protocols (unified = block, file, and object); and leverage thin and efficient technologies such as compression and deduplication to support rapidly increasing data growth. In other words, it will take the best features from today's market leading systems and apply them at scale in an always-on, shared services environment—something that previous scale-out systems have not been able to accomplish. Of course, the basic concept of scale-out storage is actually very straightforward: Instead of growing a monolithic storage system to its limits and then adding another separate storage system, scale-out storage provides a cluster of storage nodes that operate as a single entity. Adding a node to the cluster adds a predictable amount of storage capacity *and* performance, allowing users to grow storage resources incrementally as demand increases.

NetApp has been enhancing its industry-leading scale-out unified platform for many years. The most recent version of NetApp's Data ONTAP operating system introduces the capability of a data infrastructure that automatically adapts to changing workloads, provides always-on access to data, and removes barriers to scale. These innovations are delivered as part of a single operating system platform that protects customer investments and lowers total cost of ownership while directly addressing the challenge of *efficiency at scale*.

Scale-(*Way*) Out with Clustered ONTAP

Some of the key, raw attributes of Clustered ONTAP are crucial to delivering upon the promise of efficiency at scale:

- **Storage Efficiency, Intelligent Clusters, and Performance:** In mid 2012, NetApp released ONTAP 8.1.1, which provides NetApp users with a new level of achieving *Efficiency at Scale* in Clustered ONTAP environments. Some specifics:
 - For the first time, PB-scale snapshots are made possible. The value of these is in large, single namespace environments, where traditional snapshots from other vendors are simply not possible for an entire data repository. Such traditional snapshots are restricted to a single volume and those volumes themselves are restricted to relatively small capacities. In sharp contrast to this, NetApp can take PB-scale snapshots (up to 20PB to be accurate) and thereby eliminate the need for lots of little snapshots by volume. A key advantage of this approach is that it simplifies backup and replication. Overall, the PB-scale snaps mean fewer snaps to manage as well as (this is important) enough speed to be practical; this is not some arcane “World Record” ability to simply prove that something can be done (once) as a stunt. These large-scale snapshots can be used for replication, backup, and so on. The snapshot function is pretty much instantaneous,

- especially with unstructured data and even at the PB scale.
- Further sophistication comes in the form of NetApp’s “granular stop/stop capability” which means that users can pick up where they left off in the case that a scheduled dedupe does not complete in time.
 - Thin provisioning supports larger pools of shared storage for flexibility and greater savings.
 - Non-disruptive operations also get a boost by improving apparent data migration between virtual servers, basically enabling zero downtime to upgrade, and offering load balancing capability to better serve critical applications at critical times.
- **ONTAP 8.1.1:** Clustered ONTAP 8.1.1 has increased support to 24 nodes (each with up to 8TB of cache) for scale-out NAS workloads, and six nodes for block workloads (FC, FCoE and iSCSI) plus some additional optimizations and features.⁴ Historically viewed as a NAS company, NetApp is not always thought of for high-performance block workloads, but this perception will require reevaluating as its new SPECsfs and SPC benchmarks reveal.
 - **NAS (File) Performance Benchmark:** NetApp measured its Clustered ONTAP file services performance using the SPECsfs benchmark. A cluster of FAS6240 nodes demonstrated linear scaling as nodes were added with a maximum performance of over 1.5 million SPECsfs2008_nfs.v3 ops/sec using a maximum of 24 nodes. This level of performance positions Clustered ONTAP to provide the scalability and performance required to accommodate virtually any workload.
 - **SAN (Block) Performance Benchmark:** NetApp benchmarked a 6-node FAS6240 unified storage cluster for Storage Performance Council SPC-1 block performance testing running Data ONTAP 8.1.1. The results for this Clustered ONTAP SPC-1 benchmark were:
 - Productivity: ~250,000 SPC-1 I/O operations per second (IOPS)
 - Cost: \$6.69 cost per I/O per second (\$/IOPS)
 - Response: a least response time (LRT) of 0.99 millisecondsThese results demonstrate that this modular scale-out model provides a foundation for continued growth as both controller performance and node count increase over time. They are among the best SPC-1 results for enterprise disk-based systems given the low latency for the IOPS provided, and they demonstrate NetApp’s ability to enable users to *attain and sustain* the performance effectiveness element of the overall efficiency at scale challenge.

For NetApp, the path to success is being built on a unified storage architecture that addresses the majority of data center workloads, provides a consistent set of features and management tools, and scales efficiently from small and mid-sized systems to the multi-PB storage clusters required by large enterprises and cloud environments. The ability to utilize a consistent platform as you scale is significant and valuable. The NetApp model is one that not only works for mid-sized operations in and of itself, but can then also be an ideal building block approach for those organizations that scale. In other words, learn one approach and it will be usable whether you stay smaller or wind up becoming enormous, adding multiple PB of new data every year. Many midsize companies (measured, as is typical, by the number of employees) now have the kinds of amounts of data that would have been viewed as Enterprise class (as measured by the amount of digitized information they have) just a handful of years ago.

NetApp’s Advantage: Robust Storage ‘Efficiency at Scale’ Technologies

Ten core technologies currently provide NetApp users with an extensive set of tools to use in bucking the trend of coping with continual data growth, and dealing with the “tyranny of big.” These tools are the foundational elements for managing the *efficiency at scale* issues that users either already have, or will get to in time.

⁴ Note: NetApp indicates the node count is just what’s officially currently supported; there’s no hard limit.

- **Flash Cache – Performance Scaling:** NetApp Flash Cache provides an optional second-level PCIe controller-based flash cache accepting blocks as they are ejected from the system buffer cache to create a large low-latency block pool. It is currently available in either 256 GB or 512 GB modules and latency is reduced by ten times for a Flash Cache hit and 100 times for a system buffer cache hit when compared to a disk-read operation.
- **Flash Pool – Performance Scaling:** Flash Pool is a persistent, aggregate-level read and write cache. It lets users add RAID groups consisting of SSDs to a subsystem containing HDDs with the goal of delivering performance comparable to that of an SSD-only aggregate while keeping cost closer to that of an HDD-only aggregate. A relatively small number of SSDs in an aggregate is used as a persistent cache to accelerate both random reads and writes. An added advantage is that NetApp DataMotion for Volumes allows non-disruptive movement of volumes—and the associated application I/O—from one aggregate to another, enabling more focused use of Flash investment as needs change. (Where Flash Cache is tied to a controller, Flash Pool accelerates an aggregate.)
- **PB Scale Snapshot Copies – Fast Data Protection:** NetApp’s Snapshot data protection feature consumes minimal storage space because only changes made since the creation of the last Snapshot Copy are written to disk (saving as much as 50% over other snapshot approaches). As mentioned earlier, the virtually instantaneous nature of the NetApp Snapshot function is really important, as it means users can take a snapshot whenever they want without a performance hit. Some functions have been optimized for use in clusters and NetApp Snapshots can be up to 20 PB in size and performed instantly. Until NetApp introduced this capability, scale often forced users to do things differently, such as resorting to a large number of relatively small snapshot copies, despite the management challenges this produces (one such challenge being that multiple snapshots lack a consistency point, so users end up cutting and pasting, and just hoping that things did not change between the volume snaps).
- **Deduplication⁵ – Eliminate Redundant Data Blocks:** Deduplication removes redundant data blocks from volumes regardless of application or protocol. NetApp deduplication can be activated on-demand, in-line, scheduled, or when a threshold is reached. NetApp deduplication offers a unique start-stop capability which might be needed if dedupe/backups are not finished by the start of the production day. NetApp can stop, effectively bookmark the spot, and backup can proceed from that point at a later time. NetApp has over 120,000+ deduplication licenses in use and over 1 EB deployed.
- **Data Compression – Reduce Disk Footprint:** NetApp data compression reduces the size of files and provides additional space reduction when used in conjunction with NetApp deduplication. Compression is provided across primary, secondary, and archival storage, typically providing compression ratios of 2:1 or greater.
- **Deduplication and Compression Combined:** NetApp’s data deduplication and data compression are synergistic. The average UNIX or Windows disk volume contains thousands or even millions of duplicate or compressible data objects. As data is created, distributed, backed up, and archived, this data is stored multiple times unabatedly across all storage tiers. The end result is highly inefficient utilization of data storage resources.
- **Thin Provisioning – Improve Disk Utilization:** Thin provisioning can ease storage management efforts by focusing on large shared storage pools. The average UNIX or Windows disk volume can typically have 50% or more of its space unused (not utilized). Thin provisioning allows storage to keep up with a dynamic

⁵ It is worth noting that dedupe is free of charge, as are snapshot, compression, and RAID DP.

growth environment and bring applications to market faster without adding costly disk resources as quickly.

- **Thin Replication – Quickly Create Remote Backup Copies:** NetApp’s thin replication technology enables data copies that are required for protection to be created quickly and efficiently, thus reducing bandwidth limitations for branch and remote offices. Once the initial baseline data is transferred, only incremental block changes which use minimal LAN/WAN bandwidth are transferred. The result is higher performance and space efficiency when making data copies for disaster recovery and D2D backup /recovery. Combining deduplication and thin replication reduces storage requirements even more.
- **FlexClone – Instant Data Copies:** FlexClone creates temporary, writable, and instantaneous copies of data sets by creating a virtual “clone” copy of the primary data set and then storing only changes. NetApp reports that application development and test costs can be reduced by as much as 80% and enables customers to avoid making an excessive number of copies of production data.
- **RAID DP – Increase Disk Availability:** RAID-DP is a standard Data ONTAP feature using a double parity RAID 6 implementation which prevents data loss when two drives fail—potentially yielding up to a 46% capacity savings when compared to expensive mirrored disk configurations. It is also important to note that RAID–DP does not suffer the same performance hit that RAID 6 does.
- **V-Series Controller Extends Capabilities to Other Major Vendors:** The NetApp V-Series Controller brings NetApp’s efficiency solutions to other vendors’ storage systems thus preserving the investment in users’ existing storage systems and making those very systems more efficient and flexible. The V-Series provides native support for SAN and NAS protocols and scales by adding capacity to current non-NetApp storage systems or by attaching NetApp disk shelves. It should also be noted that V-Series storage can also take advantage of cluster-mode.

This is an impressive set of capabilities, and clearly the market conditions detailed extensively at the start of this paper demonstrate a clear and present need for action. But what level of results happen in the real world?

The Real World: A Case Study, and More...

PeakColo⁶

PeakColo is a Cloud Services Provider, headquartered in Denver, Colorado, that delivers a range of full-scale enterprise-class solutions, which it takes to market both directly and also through enabling VARs, SIs, and MSPs to sell their own branded versions of the PeakColo services. This model means a diverse range of demands that need to be delivered flexibly, dynamically and—by definition—*efficiently at scale*.

PeakColo chose to run on NetApp FAS3270 and 3240 storage systems with Clustered ONTAP. The results have been not only satisfying from an IT perspective, but also—and more importantly—good for the business of both PeakColo and its clients:

- Partners are able to provision storage in less than two hours.
- PeakColo is scaling at some 300% annual growth.
- It is meeting rigorous 100% availability SLAs to clients.
- It has delivered overall CAPEX and OPEX reductions via a 73% reduction in its storage requirement.

The CEO of PeakColo, Luke Norris, stated: “The flexibility of NetApp Data ONTAP 8 Cluster-Mode is invaluable to

⁶ Extracts from a published NetApp Case Study: full story can be found at <http://www.netapp.com/us/media/peak-colo.pdf>

us as a service provider, allowing us to be much more nimble than our service provider competitors. I don't know any storage operating system technology that comes remotely close to it."

More to Do: Efficiency at Scale Is a Journey, Not a Destination

NetApp is delivering on its ambitious vision, but it will need to extend its systems for even broader multiprotocol support. Although it already has all the block and file protocols covered (including newer features such as pNFS), it will need to add support for new object interfaces over time. It's important to note that NetApp possesses its own internally developed IP in addition to having acquired a powerful IP portfolio to build upon in order to offer the next-generation of storage solutions with object storage, policy management, storage service management, cross domain analysis, and advanced analytics. NetApp extended its portfolio further with its 2011 LSI Engenio acquisition for high throughput and streaming performance technologies.

The NetApp *efficiency at scale* story has really come together and is a strong one: NetApp has built a broad portfolio of solutions. Its challenge will be to deliver on the vision with even greater scale offering shared block, file, and object support; policy-enabled quality of service controls; and continued hardware advances. What's comforting is that NetApp has all the building blocks in place as well as the leadership experience to drive this, so it seems a fairly safe bet that it will continue to deliver.

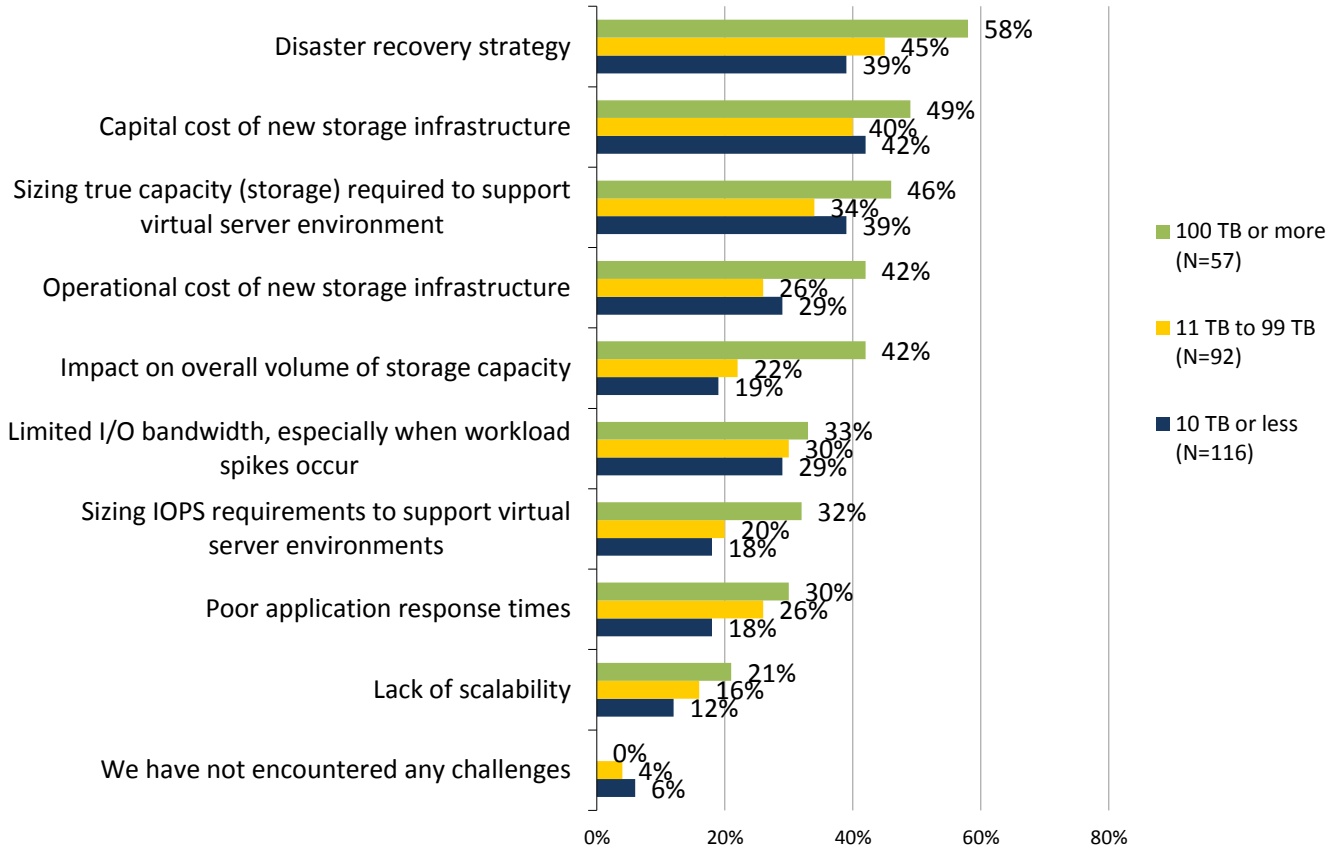
A Final Reminder....and a Recommendation

The reminder? In a world where storage devices in the 10TB range are already available for domestic use, most business users can be pretty secure in the knowledge that they will have to address the challenges of scale. And the recommendation? It is to act now...whether you are at the "gotta fix this" stage or the "we really should plan ahead stage." There seem to be certain immutable issues that make "big" difficult in IT generally and certainly within its storage component: It is, if you will, the inverse of a virtuous circle. A couple of final pieces of research serve to drive home the point that everything simply gets more challenging at scale. This means that finding a path—such as that which NetApp offers—to achieve *efficiency at scale* should be at the top of IT professionals' to-do lists. For example, Figure 4 examines the storage challenges resulting from server virtualization initiatives (chosen because it is one of the most common initiatives in IT today). The key insight is that across all the challenges, the degree to which they have been encountered is universally higher for environments with greater amounts of storage.⁷

⁷ The break point of 100TB might seem relatively small, but is explained by the fact that the capacities mentioned are only for the storage capacity used to actually support the virtual server environments

Figure 4. Storage Challenges Stemming from Server Virtualization Usage, by Total Storage Capacity Currently Used to Support Virtual Server Environment

Significant challenges related to server virtualization usage from a storage infrastructure perspective, by total storage capacity currently used to support virtual server environments. (Percent of respondents)



Source: Enterprise Strategy Group, 2012.

And, not surprisingly therefore, the desire for features and capabilities that foster storage efficiency is noticeably higher among organizations with more capacity under management. Organizations with at least 500 TB of disk storage are more likely to classify data reduction technologies, storage tiering, and unified storage as must-have features and capabilities (see Table 3). It is, if you will, a window to the necessity for efficiency at scale. Efficiency makes scalability practical in the real world—which is increasingly important as a result of rapid and rampant data growth.

Table 3. Must-have Storage Features and Capabilities, by Total Disk-based Storage System Capacity

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

Source: Enterprise Strategy Group, 2012.

The Bigger Truth

We probably all intuitively *do know* that IT and storage get harder as things get bigger. And it can be a self-perpetuating issue—for example, ESG’s research has found that annual data capacity growth rates exceeding 20% are experienced by 168% more large users (43% of those with over 500TB under management) than smaller users (only 16% of those with under 100TB). The good news is that industry leaders (whether motivated by academic logic or business necessity) are jumping in to the efficiency waters, and everyone else can learn from their endeavors...as everyone will get to the PB world sometime.

What many people probably *don’t know* is that NetApp has an impressive and comprehensive ability to provide *efficiency at scale*. Its storage efficiency offerings, combined with its enterprise feature set and ability to port to unified scalable virtualized environments, should have a direct impact on its customers’ bottom lines. Indeed, proof points show that NetApp users are successfully bucking the general trend (of spiraling costs and diminished returns) which means—for *efficiency at scale*—NetApp is taking significant steps to address the unparalleled storage challenges that lie ahead.

Combining the features of NetApp’s first-generation unified networked storage systems (that so many users know and love) with next-generation unified scale-out systems was an excellent first step for the vendor., Providing multi-tenancy and workload QoS, as NetApp is doing now, could be the core enablers taking scale-out systems to the next level and unleashing the full potential of both virtual and cloud infrastructures. If NetApp can continue to deliver on its vision, then its next decade could be just as successful and exciting as its last.



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

20 Asylum Street | Milford, MA 01757 | Tel: 508.482.0188 Fax: 508.482.0218 | www.esg-global.com