

WHITE PAPER

NetApp: Making Unified Scale-Out Storage Data Center Ready

By Terri McClure

April, 2009

Table of Contents

Table of Contents	i
Managing a Data Avalanche	1
NetApp's Unified Storage Vision	2
Storage Architectures: Scale-up versus Scale-out	3
The Economics of Scale-Out Platforms	4
Challenges with Today's Scale-Out Solutions	7
Summary	7

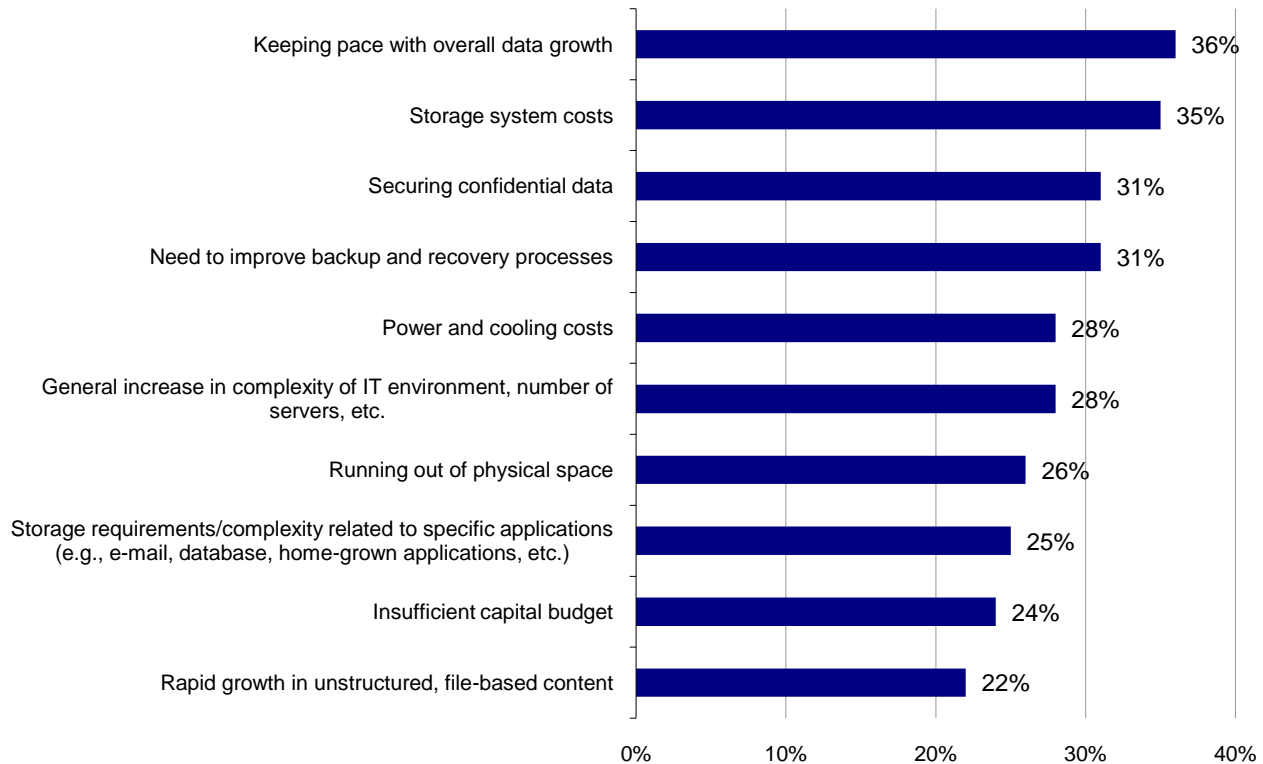
All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from 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. This ESG White Paper was developed with the assistance and funding of NetApp.

Managing a Data Avalanche

Even in a down economy, two things can be assured: data will continue to grow and IT will continue to struggle to manage that data efficiently. In fact, for some businesses, data growth is accelerating as more and more organizations use analytics to find new revenue opportunities or ways to tweak the business model and drive top and bottom line growth. So it is no surprise that ESG's research finds that keeping pace with overall data growth is the number one storage challenge faced by IT managers today (see Figure 1).

FIGURE 1. TOP 10 STORAGE CHALLENGES FACING IT MANAGERS TODAY

In general, what are your organization's greatest challenges with respect to its storage environment? (Percent of respondents, N=504, multiple responses accepted)



Source: ESG Enterprise Storage Survey, 2008

File-based data is growing fastest and will account for 70% of total archived capacity by 2012.¹ File formats are richer, creating exponential and unpredictable growth in the amount of data to be managed. Consider the growth of file data driven solely by increases in file size:

- With the digitization of media and entertainment, a two hour, high definition video can be almost 14 GB compressed (depending on compression and format) and in the multi-terabyte range uncompressed.
- Chip manufacturers render multi-terabyte files.
- Oil and gas exploration relies on 3-D models in the hundred terabyte range.
- Health care, with high definition and 4-D imaging, is creating files in the hundreds of megabytes.

This list does not even consider the impact of Web 2.0 applications, which often embed rich media and are concurrently shared by hundreds or thousands of users! Added all up, the storage implications are profound,

¹ Source: ESG Research Report, 2007 Digital Archiving Survey, November 2007.

often leading to fragmented storage deployments and “stovepiped” information infrastructures dedicated to different tiers of block or file storage. These stovepipes become a serious operational expense in terms of the number of employees required to manage the data infrastructure; excessive power, cooling, and floor space consumption; poor utilization rates—the list goes on. Unified storage is one way IT can get its arms around managing massive data growth while reducing operational costs.

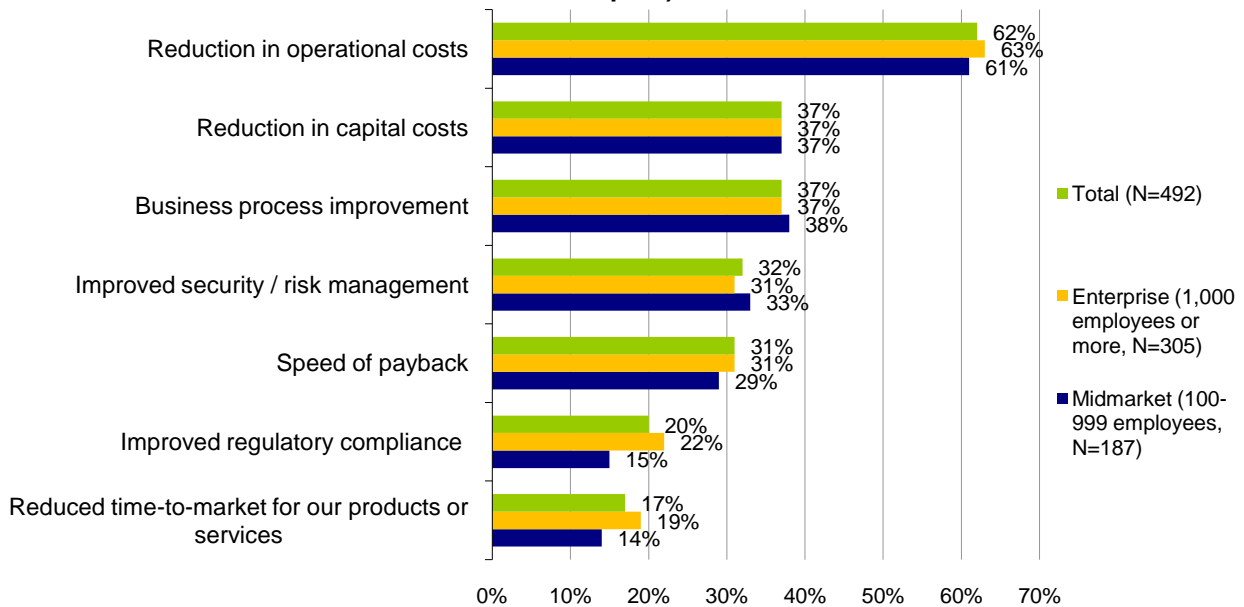
NetApp’s Unified Storage Vision

When most storage vendors talk about unified storage, they mean the ability to support multiple data types (block and file) and network protocols (Fibre Channel, iSCSI, and/or FCoE) within the same array. For NetApp, it means more than just having platforms that support multiple data types and data access protocols. NetApp includes common components and management across storage tiers, protection, disaster recovery processes, provisioning, you name it; NetApp’s version of a unified storage platform includes a soup-to-nuts standardization of the storage environment. This is an important point—from an operational standpoint, it means only training administrators on a single management interface—no matter what tier of storage, from primary to archive, block to file. It also means IT can standardize on data protection and disaster recovery processes across tiers as required.

In the current economic climate, cost reduction is king. ESG recently conducted an in-depth survey of 492 senior IT professionals concerning their organizations’ IT spending plans and priorities over the next 12-24 months. For both midmarket and enterprise IT professionals, reducing operational costs far outweighed any other criteria for justifying IT purchases over the next twelve to twenty-four months (see Figure 2).²

FIGURE 2. REDUCTION IN OPERATIONAL COSTS IS KEY TO JUSTIFYING IT SPENDING

Which of the following considerations do you believe will be most important in justifying IT investments to your organization’s business management team over the next 12-24 months? (Percent of respondents, multiple responses accepted)



Source: ESG Research Report: 2009 Data Center Spending Intentions Survey, March 2009

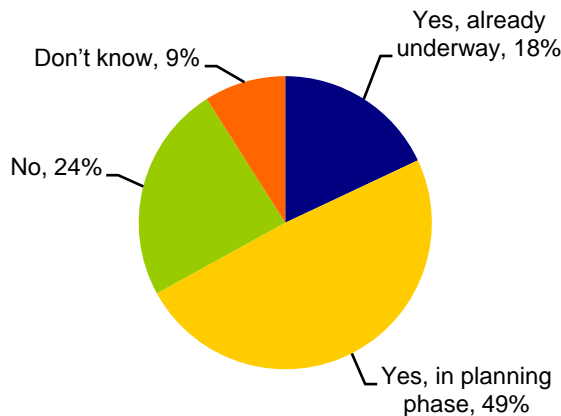
Users recognize the benefit of unified storage in the broad definition of a platform that supports multiple protocols. In a recent survey of 338 IT professionals with storage buying responsibility, 67% are either deploying

² Source: ESG Research Report, 2009 Data Center Spending Intentions Survey, March 2009.

or planning to deploy unified storage. The days of storage stovepipes in the data center are over—management is just too hard and consumes too many cycles.

FIGURE 3. UNIFIED STORAGE ADOPTION

Does your organization have any plans to consolidate NAS and SAN storage resources into a unified storage architecture that supports both file-based/NAS and block-based/SAN storage? (Percent of respondents, N=338)



Source: ESG Enterprise Storage Survey, 2008

NetApp has a solid unified storage foundation with its Data ONTAP operating environment and is committed to taking Data ONTAP unified storage to the next level. Scale-out capabilities are the next piece of the puzzle that enables users to continue consolidation. Just as it simplified NAS management more than a decade ago, NetApp's current vision is to extend that experience to unify and simplify managing the entire storage ecosystem. This means providing users with a simple dashboard, allowing them to specify storage service levels while the back-end operating environment assigns resources based on pre-defined policies. The services encompass adjusting processor and disk profiles for block, file, scale-up, and scale-out. The goal is policy-based management where application owners can specify service levels for their own applications without requiring storage administrator expertise.

Consider the power of managing, under a single umbrella with standardized processes, a storage ecosystem that supports file- and block-based data access, scales to massive capacity, and supports scale-up and scale-out profiles. Layer on market-proven value-add software and, because scale-out cluster environments essentially virtualize the underlying storage, you have a truly liquid environment that can evolve with the business and enable business agility. It is a powerful vision.

Storage Architectures: Scale-up versus Scale-out

To understand why scale-out is a key addition to the unified storage platform, it is important to understand what it is and how it is differentiated from scale-up architectures. Scale-out, the ability to independently scale and tune bandwidth, processing, and storage capacity on the fly—all while managing the storage system as a single system image—is designed to overcome the management and performance challenges of handling large amounts of data. It has seen early adoption primarily in the NAS space thanks to a need to support ever-increasing file sizes and the speed of file data growth. But scale-out is also starting to emerge as a SAN requirement. Since scale-out architectures virtualize the storage environment, users are insulated from underlying hardware changes. This type of abstraction is crucial in enabling a nonstop, always-on enterprise infrastructure and significantly increases reliability, availability, and serviceability. Users are looking at new technologies that will allow them to scale to previously unheard-of capacities, availability, and throughput using scale-out architectures. These storage architectures differ significantly from the scale-up storage architectures that were developed to meet distributed computing needs.

That's not to say that scale-up is bad; scale-up storage architectures continue to be suitable for the I/O-intensive transactional environments they were developed to support. But scale-up architectures have limited throughput and bandwidth and are not designed for environments handling very large files or requiring lots of bandwidth for throughput—files that are single-writer and multi-reader, with high concurrency of access or very large sequential reads and writes. Scale-up storage is just what it sounds like—it is designed to be monolithic, with lots of storage sitting behind one or two storage controllers, and it is designed to scale into the multi-TB range behind those controllers. Once the limit on storage is hit, a new system, with a new file system to manage, is installed. Scale-up systems have no way to scale bandwidth without some price penalty. High sequential performance in today's scale-up systems is often scaled by adding a storage rack and more spindles to increase throughput and reduce latency (and, as a byproduct, reduce storage utilization)—an expensive proposition.

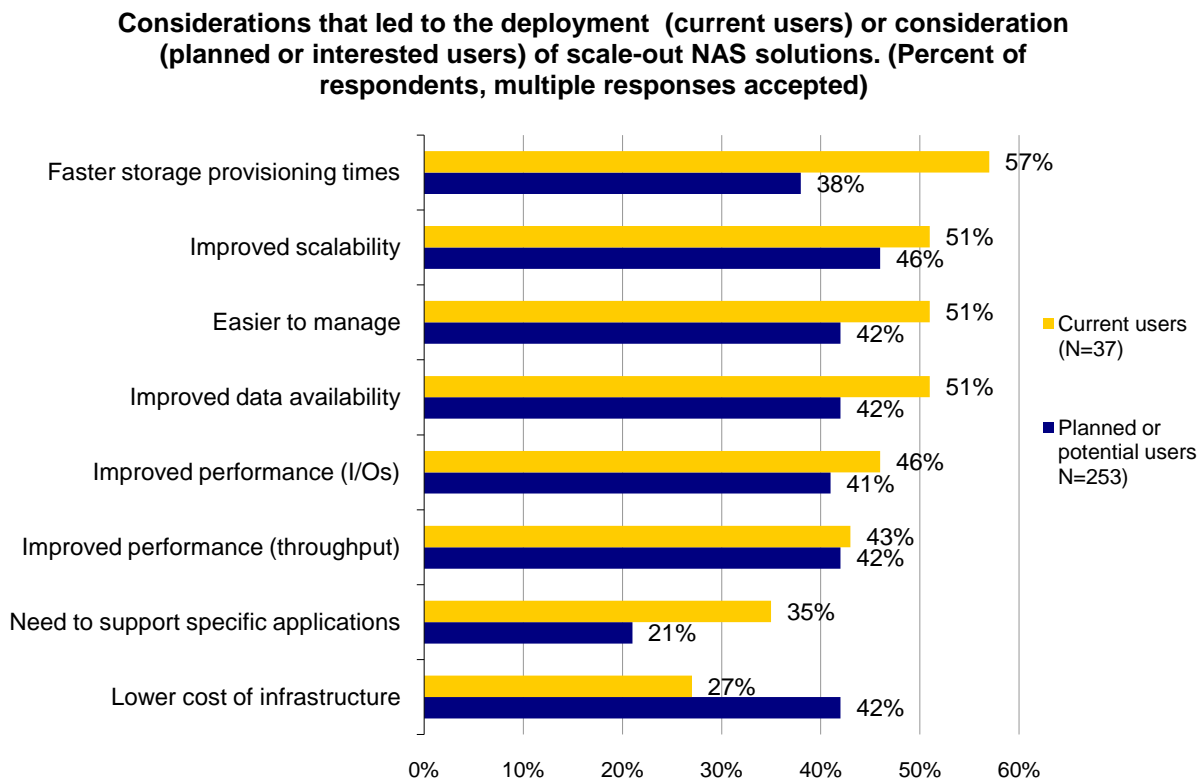
Scale-out NAS is not a new concept. Users in high performance computing (HPC) and the media and entertainment market were early adopters because of its bandwidth. As commercial enterprises adopt Web 2.0 applications and commercial HPC as file data continues to grow at an accelerating pace, scale-out technology is becoming a core data center requirement. NetApp has been shipping the scale-out Data ONTAP GX system for the past several years. But NetApp's GX was just the first step in bringing scale-out solutions to market—their vision is to bring scale-out into the unified storage fold, therefore providing scale-out SAN as well as scale-out NAS. The result should be a multi-protocol storage infrastructure that addresses a diverse set of workloads across enterprise IT and emerging cloud based application services.

The Economics of Scale-Out Platforms

Scale-out systems scale at a very granular level, adding processing power or capacity when and as needed; users can start with as few as two nodes and scale up from there. Because scale-out systems virtualize the underlying storage hardware and insulate users from any changes in the hardware, capacity can be added on the fly without taking the system offline.

Scaling granularly provides a price/performance advantage as it allows users to start small and scale where and as needed. And since scale-out systems scale into the multi-petabyte range in which multiple nodes are managed as a single system image, the systems can meet most users' needs without paying the management penalty associated with deploying tens or hundreds of scale-up systems. Additionally, with most scale-out systems, many of the low level storage management tasks are automated, such as expanding the file system when new physical capacity is added and load balancing performance across processors to further reduce management costs.

For users evaluating new solutions, initial cost has become a higher priority than the value over time, though scale-out systems provide cost advantages in both areas. ESG recently conducted a survey of 504 North American and Western European IT professionals to assess data storage environments, including the adoption of scale-out NAS. Market drivers for early adopters included faster provisioning, improved scalability and performance, easier management, and the need to support specific, fast-growing applications. Lower cost of infrastructure was literally *last* on the list of buying criteria. However, planned and potential users have vaulted lower cost into the top tier of purchasing criteria—second only to improved scalability, which is the crux of the technology (see Figure 4).

FIGURE 4. SCALE-OUT NAS ADOPTION DRIVERS

Source: ESG Research Brief, *Scale-Out NAS Adoption and Market Drivers*, February, 2009

Scale-out architectures have a number of cost advantages over scale-up solutions, ranging from start up costs to managing technology refreshes—and many steps in between. Scale-out carries a lower overall cost compared to scale-up systems for a number of reasons:

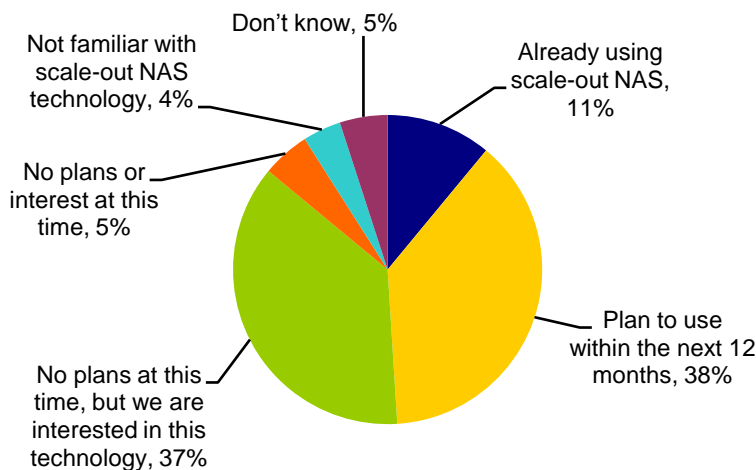
- **Low entry cost:** The entry cost for scale-out systems varies depending on the minimum configurations supported. Most systems start as small as two nodes and scale-up from there. With clustered scale-out systems, you can add resources and scale as needed.
- **Just-in-time scalability:** As previously stated, because of the modular nature of scale-out systems, there is no need to buy (and power or cool) frames, power supplies, and mostly empty cabinets in advance of storage capacity.
- **Riding the commodity curve:** Scale-out systems typically use low-cost, high capacity, commodity SATA disk drives. Some systems use combinations of Fibre Channel and SATA to tier storage within the system, reducing the overall number of drives required. Either way, users are able to take advantage of falling disk prices thanks to the just-in-time scalability of scale-out storage. The same cost advantages are typically found on the storage controller, where commodity processors are used. Riding the Intel and high capacity disk commodity curve can add up to a significant cost savings, especially in large scale environments.
- **Higher utilization rates:** Better utilization means deferred purchases of new capacity. Because all of the processor heads in scale-out systems can address the entire pool of usable capacity in the cluster, no capacity is locked away behind underutilized processors—a common problem in scale-up systems. It is not unusual to see utilization rates of 30% or less in scale-up systems and 60% or more in scale-out systems.

- **Reduced change management planning cycles:** When every endpoint device is a content capture device, conventional three or six month change management planning cycles are no longer effective. Requirements are unpredictable and time-to-provision is more important than ever. The modularity and scalability of scale-out storage allows for extremely fast provisioning. Most systems are plug-and-play; add a storage or processor node and the system self-discovers and expands the file system or incorporates it into load balancing algorithms on the fly. There is typically no disruption of service, nor is there a requirement to plan data layouts, create LUNS, or migrate data.
- **Non-disruptive technology refresh.** With most scale-out systems, the process of managing technology refreshes is faster and easier than with scale-up NAS because the global namespace maps logical mount points to physical mount points in a virtualized manner, allowing back-end technology changes to be made with little or no disruption to client access.
- **Ability to scale capacity without scaling headcount.** Essentially, it should be just as easy to manage a clustered storage system with 100 nodes as it is to manage one with two nodes. Scale-out file storage systems enable this through a global namespace, which provides a single point of management for massive amounts of file data.
- **Automated low-level functions:** Most scale-out storage systems support deep levels of policy-based self management and healing. Scale-out systems typically absorb new processor, bandwidth, and storage capacity, then automatically re-balance and optimize across the newly added resources—with little or no human intervention.

Based on the compelling economic benefits of deploying scale-out solutions, it's no surprise that recent ESG research indicates that users are applying scale-out NAS systems to new use cases. While most scale-out systems are tuned to perform well for high bandwidth applications, some can also support the smaller transaction-oriented file serving requirements of today's distributed computing environments. In fact, 43% of scale-out NAS users recently surveyed by ESG indicated that the technology is used to support database and OLTP transactions. Further proof that scale-out NAS is increasing its footprint in the general storage space is that even though only 11% of those surveyed indicate they use scale-out NAS systems today, 40% indicated they plan to deploy it within the next 12 months. And while another 37% have no immediate plans to deploy scale-out NAS solutions, they are investigating the technology (see Figure 5).

FIGURE 5. SCALE-OUT NAS ADOPTION

Please describe your organization's usage of or interest in scale-out NAS systems. (Percent of respondents, N=338)



Source: ESG Research Brief, *Scale-Out NAS Adoption and Market Drivers*, February, 2009

Thanks to the previous points, more and more enterprises are taking a serious look at scale-out storage—clustered scale-out solutions are going mainstream. But commercial enterprises are not just interested in the increased bandwidth scale-out solutions bring to the table. Users are expanding use cases for scale-out thanks to the higher scalability and manageability of these systems. Clearly, scale-out makes economic sense. The challenge now is finding scale-out solutions with the enterprise class features already so familiar in the data center. These would include not only the use of scale-out to create a global namespace for file-based storage (NAS), but also the ability to scale-out the enterprise SAN infrastructure.

Challenges with Today's Scale-Out Solutions

Most commercially available scale-out solutions are designed to solve a single problem in environments that are compute intensive, such as batch processing, or interactive & collaborative, such as product development—forcing users to continue to buy specialty solutions tuned to each need. Many came from HPC technical and niche markets and lack enterprise-class features that were not core requirements in those markets, such as remote mirroring and snapshots. And almost all solely support file or block storage, not both.

With death, taxes, and data growth being the only sure things in life, users cannot afford to continue installing specialty arrays for each and every compute environment. Users need to have tiers within arrays for scale-up, scale-out, high-performance, and bulk storage—managed as a single system image. And they need to be able to protect data in line with its value using capabilities like snapshot and local and remote mirroring.

Summary

Facing the stark reality of a prolonged economic slowdown, users are looking for a number of things from their storage vendors. They want vendors they can trust, market proven solutions, and products that offer real cost savings. New rich media content is being created for everything from training to marketing and becoming a mandatory component of everyday business. Whether it's blogs or video, content is easier than ever to create—and management will become harder than ever without significant changes. Keeping up with data growth driven by the ubiquity of content capture devices (i.e., every endpoint device), new types of applications, and richer media types requires a new approach to keeping storage costs in check. Scale-out is that approach. Enterprises that deploy scale-out solutions can get more value, dollar-for-dollar, from their infrastructure investments. Scale-out has a compelling value proposition; its lower infrastructure costs, power efficiency, and management efficiencies at scale should put scale-out solutions on the short list for anyone deploying new capacity.

If you plan to put scale-out in the data center, carefully evaluate your options. The market is not fully mature and many of the offerings available today do not have the enterprise-class features you are used to from your data center NAS and SAN vendors. But scale-out architectures can help users respond to changing business demands quickly and economically and are starting to be offered by trusted data center storage vendors like NetApp.

NetApp's vision—to offer a truly unified storage experience that lets users scale up or out, for block or file storage, and maintain a single system image—is a compelling one. NetApp's strength continues to be its unified storage platform, supporting both block- and file-based storage—the efficiency benefits of scale-out apply as much on the block side of the equation as on the file side. While most vendors are offering either NAS or SAN and stovepiped solutions customized to meet needs for each tier of storage, NetApp is continuing its focus on bringing new value to unified storage—scale-out is a natural extension. NetApp's implementation of a unified storage approach to scale-out should provide a solid, scalable foundation for emerging cloud based infrastructure and services. NetApp has been shipping the GX platform for several years now, which is a good proof point, but was also only a first step. Expect NetApp to raise the bar over time, continually building scale-out features on the Data ONTAP unified foundation.



20 Asylum Street
Milford, MA 01757
Tel: 508-482-0188
Fax: 508-482-0218

www.enterprisestrategygroup.com