

# Nexsan and FalconStor Team for High Performance, Operationally Efficient Disk-based Backup

**Date:** August, 2009

**Author:** Brian Babineau, Senior Consulting Analyst, and Lauren Whitehouse, Senior Analyst

**Abstract:** Disk-based backup, deduplication, and MAID are three of the most talked about technologies in the storage industry—each is so popular that organizations often purchase them individually just to see the capabilities they can bring to their environments. Now, all three are available in Nexsan's DeDupe SG—an integrated system from Nexsan and FalconStor that promises to optimize almost any backup environment.

## Overview

By now, most organizations are not wondering *if* they should evaluate disk-based backup to expedite data protection processes. Rather, they are trying to figure out what solution is best for them—a difficult task given the amount of data requiring protection continues to increase as IT budgets being are cut. The evaluation criteria list for disk-based backup solutions is likely to change; performance is no longer enough. Now, a solution must also be able to cut data center expenses as nearly two-thirds of those companies recently surveyed by ESG cite reduction in operating costs as the best way to justify an IT investment over the next two years.<sup>1</sup> Translated—organizations are more likely to evaluate disk-based backup targets that are easy to implement and manage, that reduce the amount of capacity needed to store a backup, and that minimize power consumption while being able to actually complete the backups within the designated window.

The emergence of new purchasing requirements is a likely catalyst that will spur customers to evaluate the recently introduced Nexsan DeDupe SG—the culmination of a joint venture between Nexsan Technologies and FalconStor Software. The Nexsan DeDupe SG solution family combines Nexsan's storage systems with FalconStor's File-interface Deduplication System (FDS) deduplication software in a fully-integrated system. Available through any of Nexsan's 300-plus worldwide channel partners, the Nexsan DeDupe SG family consists of six models which scale from 4 TB to 52 TB useable capacity (the amount of capacity that is available to an application after RAID formatting and configuration for 'hot spares') to address the various backup capacity requirements of small to mid-size enterprises.

## Disk-Based Backup Refresher

Despite current economic conditions, ESG estimates that corporate data will increase 25% in 2009 after growing at two to three times that rate for much of the decade.<sup>2</sup> IT must find ways to adequately protect this new information within the allocated backup window—with flat to decreasing budgets.

Introducing disk enables customers to shrink backup times as it is faster than tape-based backup, but this is not all there is to the decision; the acquisition cost of disk is more than tape. Companies that cannot afford the upfront capital investment tend to stick with tape and deal with longer backup windows. Buying less capacity is the easiest way to lower the cost of disk and make the investment profile more comparable to tape. And, the easiest way to achieve this is by deduplicating redundant data during the backup process, minimizing the amount of data to be stored, which in turn minimizes the amount of disk that has to be purchased for data protection purposes.

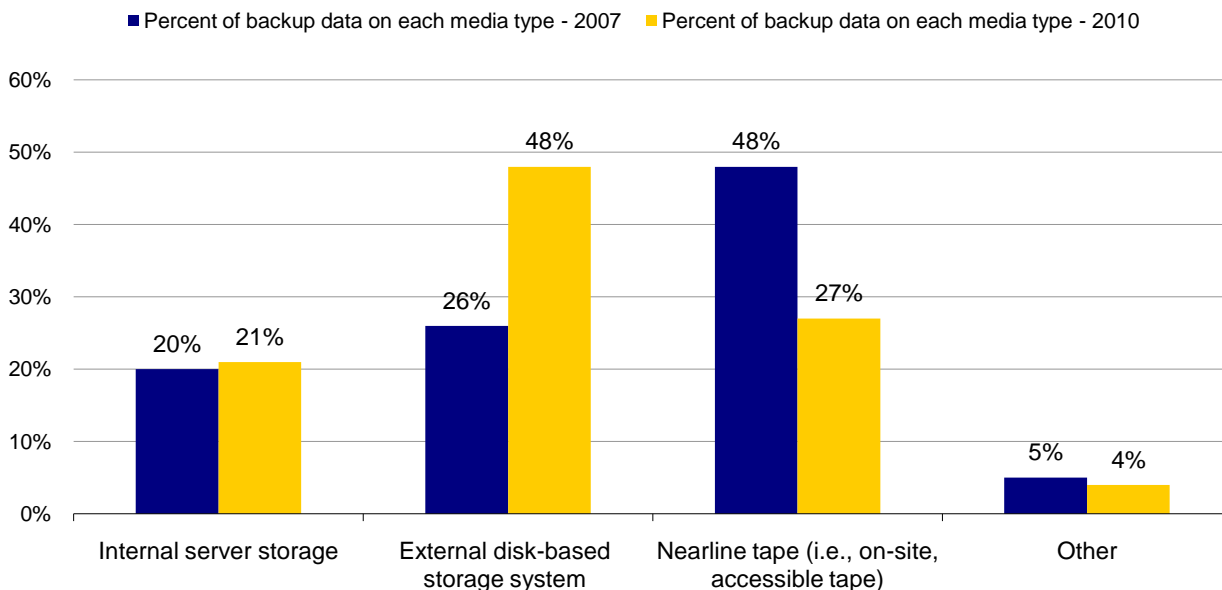
Deduplication solutions are driving the transition from tape- to disk-based backup. According to an ESG survey of over 350 organizations, the average respondent estimated that 48% of their onsite backups would be stored on external disk by 2010—up from 26% in 2007 with much of the capacity being shifted from tape (see Figure 1).

<sup>1</sup> Source: ESG Research Report, *2009 Data Center Spending Intentions Survey*, March 2009.

<sup>2</sup> Source: ESG Research Report, *ESG 2008 Enterprise Storage Systems Survey*, November 2008.

**FIGURE 1. EXPECTED INCREASE IN ONSITE EXTERNAL DISK SECONDARY CAPACITY BY 2010**

**Approximately what percentage of your organization's total onsite backup data is currently stored on each of the following storage media types? Please also indicate what you expect these percentages to be in 2010? (N=364)**



Source: ESG Research Report, Data Protection Market Trends, January 2008

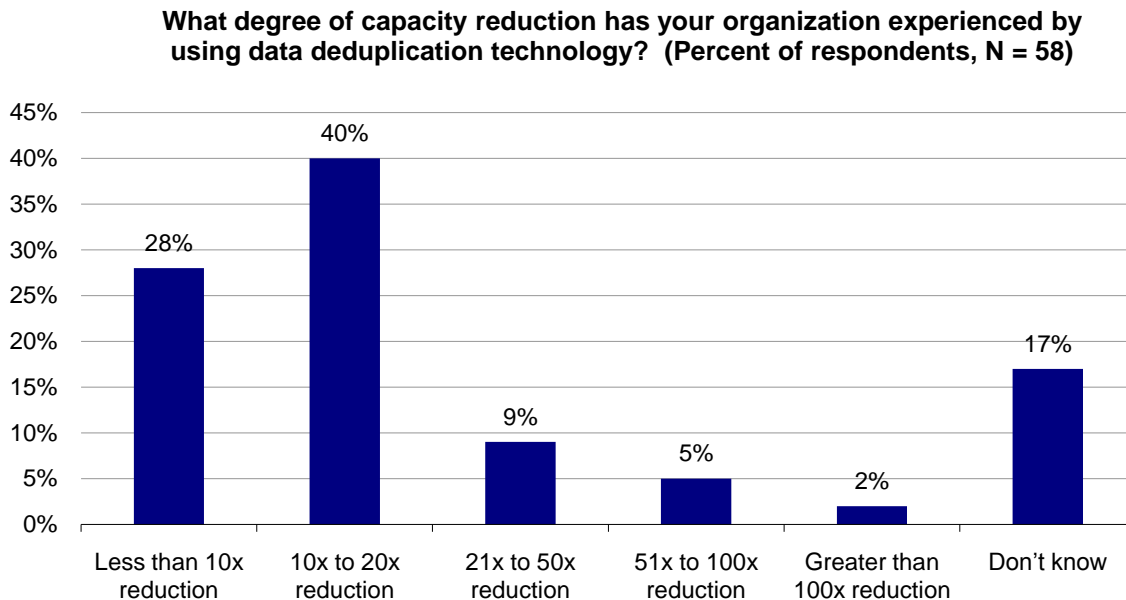
Given the complementary nature of the technologies—deduplication solutions and reasonably performing, dense disk systems—organizations are going to evaluate both. There is plenty of each type of offering in the marketplace; that’s not the issue. Buying them separately requires integration. Organizations could choose from a few pre-integrated systems, but many are complex to implement and manage or cannot scale to meet corporate data growth.

**Nexsan DeDupe SG: Flexibility with Simplicity**

A Nexsan DeDupe SG is made up of two components—a server running FalconStor’s FDS and Nexsan storage. The only choice customers have to make is which of the six Nexsan DeDupe SG models best addresses their capacity and performance requirements.

During selection, customers must take into account the data reduction ratios for their various information sources. This is a function of how often backups are run, the type of backup job for a specific application (i.e., incremental, full, or differential), and retention settings, as well as the type of information contained in a data set. For example, server virtualization images, such as VMware’s VMDK files, contain much of the same information and if these are constantly backed up, they will have a much higher data reduction ratio than files which do not contain a significant amount of redundant content. In discussions with several organizations currently using deduplication, ESG has seen data reduction ratios range from 4:1 to 50:1 (see Figure 2).

**FIGURE 2. CAPACITY REDUCTION RESULTING FROM DATA DEDUPLICATION TECHNOLOGIES**



Source: ESG Research Report, Data Protection Market Trends, January 2008

Customers must also consider their performance needs when selecting the appropriate Nexsan DeDupe solution as it is designed to scale in this dimension as well. The bigger capacity systems, fitted with more RAM, can ingest data faster. The faster the ingest process, the faster backups complete. This enables customers to consolidate smaller backup environments on a single Nexsan DeDupe SG. Customers may also be attracted to the larger systems to achieve higher data reductions ratios—more data being sent to the same location increases the likelihood that the same bytes are being copied, creating more opportunities for deduplication.

Customers will enjoy the flexibility offered in the various Nexsan DeDupe SG capacity models along with the simplicity of purchasing a solution. They will also be pleased to learn that Nexsan is also handling customer service/support and has worked out how escalations with FalconStor will be handled as part of the joint venture.

**How it Works**

When a customer orders a Nexsan DeDupe SG, they get a system that includes a deduplication server powered by FalconStor’s FDS and a Nexsan storage system (either a SATABeast or SATABoy). The customer simply racks the system, plugs the server into the storage, and powers them on. Upon startup, customers will be presented a configuration wizard to set up the storage volumes and associated file systems to their liking.

The system connects to any Ethernet network via four or eight (depending on the model) 1GbE ports and presents a standard file system interface—NFS or CIFS—to any device that customers want to back up. In some cases, the Nexsan DeDupe SG will be the target of a backup application server; in other instances, customer may want to connect directly to a primary file server that needs to be protected. In the latter scenario, customers can execute ‘data dumps’ using scripts or application specific tools such as Oracle RMAN to copy files directly to the Nexsan DeDupe SG.

Data received by the Nexsan DeDupe SG is analyzed by FalconStor FDS—every block of data written is compared against the central hashing index. If the block is unique, it is stored in the deduplication repository and added to the index; if the block already exists in the deduplication repository, FDS then points to the existing block and deletes the newly received, redundant data. Running the deduplication process after the data is written on disk ensures that the backups are complete before any redundant data is actually removed.

With the Nexsan DeDupe SG, customers have the flexibility to schedule when the deduplication process occurs—some may run it concurrently (as data is being backed up it is also deduplicated) or after all backup jobs are

complete. In addition, the deduplication policy engine enables customers to exclude some data streams from the deduplication process, which speeds up the overall backup operation because CPU resources are shifted to information sources that are more likely to have redundant data.

Customers can retrieve files directly from the Nexsan DeDupe SG or go through the backup application that copied the data. Because all data is stored in a file system, there is no restore process necessary—as there is when data is stored in tape format—before data can be accessed.

## Performance Matters

The primary reason why a company introduces disk into the data protection process is to reduce the time it takes to complete backup and restore operations. Adding deduplication to a disk-based backup solution should not hamper this performance gain—companies still need to create copies of data within an allotted timeframe. Both the deduplication operations (ability to process more data) and storage device (ability to handle the data being sent by the deduplication engine) must scale as more data is backed up to the system. Scaling is extremely important as the amount of data processed and initially stored is much larger than the data that is ultimately saved (after the deduplication process occurs).

The entry level Nexsan DeDupe system, which has 4 TB of usable capacity, can ingest data at 150 MB/s, while the 52 TB solution can receive data at 750 MB/s through standard CIFS and NFS protocols. The middle four members of the Nexsan DeDupe SG family scale at a similar capacity to ingest ratio. Although ESG Lab has not tested a Nexsan DeDupe SG and every customer's environment is unique, the published specifications compare favorably with other disk-based backup systems that offer deduplication. What's even more impressive is that Nexsan DeDupe SG specifications are achievable with any backup software application, application-specific backup tool, or script. Some deduplication system benchmarks can only be attained with backup software options, which are optimized to write to disk such as Symantec's NetBackup OST. While these backup software options can enhance a disk-based backup environment, they are not free and only work with certain hardware and server configurations (i.e.: a certain host operating system attached to specific disk systems)—the performance gains may not be enough to justify the investment.

Nexsan and FalconStor are also evaluating ways to allow customers to further improve performance. For example, they may test and qualify 10GbE NICS for faster front end connectivity. These efforts, along with the initial performance specifications, should prove to customers that Nexsan and FalconStor are determined to make disk-based backups with deduplication complete as quickly as possible.

## Speed with Green

If a customer completes backups in a given portion of the day, a significant amount of energy is wasted as the target storage system remains online, waiting for the next data feed or request. Disk drives continue to spin and must be cooled, even if data is not being written. This presents a great opportunity for IT to cut energy costs—an opportunity that many organizations plan to take advantage of as nearly 42% of organizations recently surveyed by ESG said that they planned to maintain or increase their investment in more power-efficient storage hardware in 2009.<sup>3</sup>

Customers can lower power consumption and support green initiatives with the Nexsan DeDupe SG as all the systems support AutoMAID, a process that idles or shuts down inactive drives. By idling or shutting down inactive drives, the Nexsan DeDupe SG consumes less energy (the drives do not draw any power) and reduces the amount of energy needed to cool the solution (because the drives aren't spinning and generating any heat). Customers do not have to install any software on (backup application) servers or other hosts connected to the Nexsan DeDupe SG to run AutoMAID.

The best part about AutoMAID is that it is not an 'all or nothing' option. Customers configure AutoMAID to balance energy savings with drive response times (this is what the SG in the product name stands for: "Speed with Green"). For example, a level-one AutoMAID setting ensures sub-second drive response times while delivering 15-20% energy savings as the drives are still powered and can spin up immediately. When AutoMAID is set at level-two, customers can expect a 15 second response time and achieve a 35-45% reduction in energy

---

<sup>3</sup> Source: ESG Research Report, *2009 Data Center Spending Intentions Survey*, March 2009.

consumption. Level-3 delivers 60-70% energy savings and drives will respond to a request in 35-40 seconds. All of the aforementioned response times are for the first IO request only—all subsequent IO requests are handled at full speed as the drives are spinning at their maximum rates. This flexibility enables customers to further optimize their backup environments by reducing energy usage without impacting the ability to access data in a timeframe acceptable to the business.

### Ideal for Mid-Size Enterprises

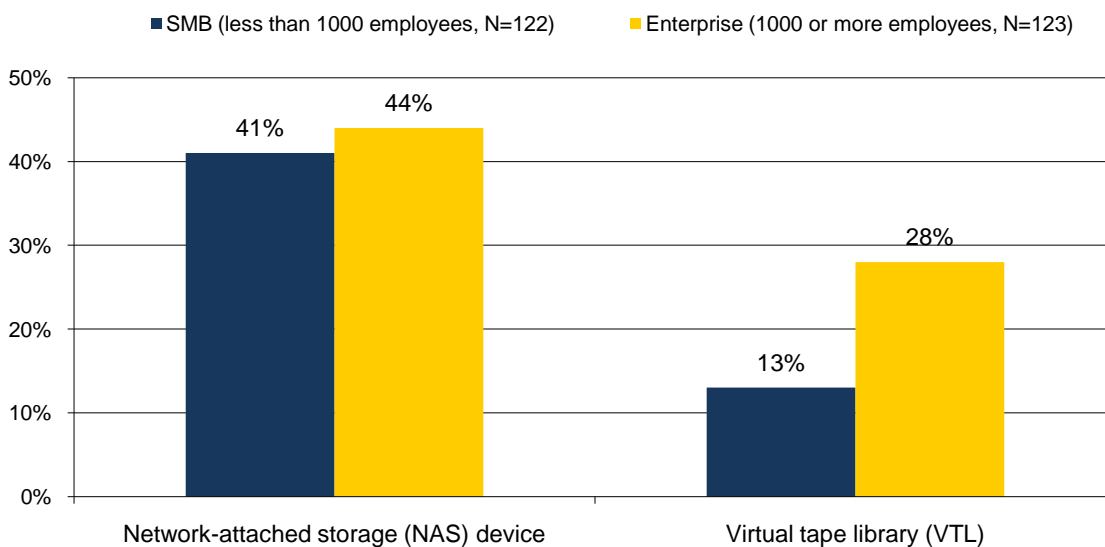
Small to mid-size enterprises, and even departments of large enterprises, rarely have the resources to buy ‘best of breed’ or disparate technologies and integrate them. They do not want to comb through complicated vendor price books to configure a backup solution that seemingly meets their needs. And they certainly do not have the time to implement and manage a complex disk-based backup environment—they are trying to eliminate this complexity by replacing tape.

There are multiple reasons why Nexsan’s DeDupe SG is perfect for mid-size enterprises, the most notable being it is one integrated product that comes in six models varying only in capacity. The ‘out of the box’ setup requirements involve plugging in a few cables and completing a short configuration wizard. Service and support of the solution is straightforward, as is finding someone who offers the solution—it’s only available through Nexsan’s worldwide network of resellers.

Another reason why the Nexsan DeDupe SG will work well in small to mid-size enterprise environments is its connection method: The Nexsan DeDupe SG is essentially a Network Attached Storage (NAS) device that installs on an Ethernet Network and presents an NFS or CIFS file system interface to a server that it is going to protect. Most companies have IP expertise in house and know how to manage systems that communicate via standard file system protocols. In contrast, many other disk-based solutions connect to a target via a Virtual Tape Library (VTL) interface, which requires a Fibre Channel (FC) Storage Area Network to be in place. Organizations then have to configure the libraries and tape drives within the VTL as well as purchase and run the FC network. As a result, smaller organizations prefer the simpler NAS disk-based backup option while larger enterprises with significant tape installations and specialized IT resources favor a VTL solution (see Figure 3).

**FIGURE 3. ENTERPRISE AND SMB DISK-BASED BACKUP TARGET PREFERENCES**

**What type(s) of onsite disk storage systems does your present location use as a backup target? (Percent of respondents, multiple responses accepted)**



Source: ESG Research Report, Data Protection Market Trends, January 2008

## The Bottom Line

There are plenty of disk-based backup and deduplication products available today. There are even integrated solutions that combine these two capabilities being offered—many of which can help customers optimize their backup environments. However, given today's realities of growing data, shrinking budgets, and smaller IT staffs, making backups run faster is not enough. Organizations need fully-integrated disk-based backup solutions with deduplication that are also easy to buy, reduce energy costs, and simple to install and manage—all of which reduce operating costs

Logic dictates that customers will initially evaluate the Nexsan DeDupe SG solution because of its performance metrics (as it certainly helps backups execute faster) and its ability to reduce the overall amount of data being stored for data protection purposes. But, when it comes to making an investment, it is likely the customers will be drawn to the Nexsan DeDupe SG's standard file system interface, flexible AutoMAID settings, and configuration wizard as these deliver compelling operational savings and efficiencies that drive an immediate ROI.