

## ExaGrid: An Emerging Leader A Compelling Disk Backup Story

**Date:** August, 2007

**Author:** Heidi Biggar, Analyst

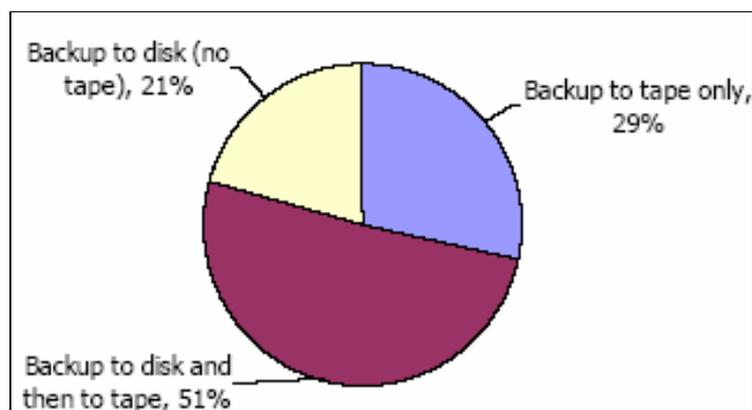
**Abstract:** ExaGrid is an expert on the topic of data reduction as well as a viable competitor in the midrange disk-based backup market. The company has consistently gained presence with customers by providing products that just plain work.

### Disk-Based Backup Matures

What a difference a year makes—not just for ExaGrid, but also for the disk-based backup market in general. Over the last 12 months, ExaGrid has watched the sales and implementation of its disk-based solutions grow significantly while the industry has witnessed the maturation of disk-based solutions as a whole. Disk-based data protection is still an emerging market, but user comfort level with—and implementation of—these types of solutions continues to increase. Disk-based backup products have become more widely available and, importantly, are being used by more and more organizations—big and small. In fact, in the mid- and small enterprise markets where ExaGrid predominantly plays, ESG Research finds that users are more likely to replace entire tape-based data protection infrastructures with disk than they are in larger organizations where tape investments are significantly larger.

ESG has been an advocate of disk-based backup for several years. Why? The potential benefits speak for themselves: improved backup and recovery performance, easier management, better reliability, etc. Backing up to and recovering from disk-based platforms is not only faster than tape, but it is also more efficient (from a management standpoint) and more reliable, largely because there is a lot less human handling required. And with the advent of capacity optimization technologies such as compression, delta differencing and data de-duplication, the appeal of disk-based backup is that much greater. Reduced capacity requirements, longer retention periods and new and/or expanded remote backup options are all byproducts of data de-duplication.

**Figure One: Current Data Backup Process**



As a barometer of user interest in and adoption of disk-based technologies, ESG Research surveyed 228 IT professionals, asking them whether they use tape or disk in their current data backup process. 29% of respondents use tape exclusively, another 21% were disk-only shops and 51% of users surveyed employ

some combination of both disk and tape.<sup>1</sup> This means that 72% of respondents use disk in some aspect of their backup process, which is strong evidence that disk is quickly becoming the go-to backup platform (see Figure One).

Clearly, user data protection behaviors are changing. Organizations are storing months of data on disk-based backup systems as insurance to recover data if needed. This is very different from how they use tape, which they rotate weekly or even daily. Recovering a single file from disk can be hundreds of times faster than recovering from tape. But the real performance benefits come from being able to instantly restore a file from disk versus having to search for tapes that are no longer in the tape library. The performance challenges with tape increase when related files or database tables span multiple tapes. ESG has spoken with numerous customers who tell the same story of trying to restore data that spans multiple tapes and has been vaulted off-site—a process that takes days or even weeks to complete.

Organizations are keeping data retained on disk for months—solely for the purpose of rapid recovery. However, there is real cost associated with storing months of backup data on disk-based storage systems. And, again, that is where capacity optimization fits in. The amount of storage can grow quickly, reaching several multiples in capacity compared to the original primary data in a very short period of time. Capacity efficient technologies can change the economics of disk-based data protection by significantly reducing the amount of redundant data that is backed up both locally and over the WAN.

### ExaGrid's Approach

ExaGrid is a provider of disk-based backup storage systems. Administrators simply back up to an ExaGrid file system either through the NFS or CIFS protocols. This process is simple and non-disruptive. ExaGrid provides more than just a disk target. Because ExaGrid performs byte-level de-duplication, it can reduce significantly the amount of physical capacity required to store backup data. Assuming 20:1 reduction, companies can store 20 TB of backup data on 1 TB of physical capacity. This level of capacity optimization drastically changes the economics of storage by driving down the effective cost per GB stored by the same factor. For further efficiency, ExaGrid also compresses the data.

To put this in perspective, consider the following hypothetical comparison illustrated in Figure Two. The traditional customer (left) starts out backing up 1 TB of data. This customer creates 20 GB of new data each week and does nightly full backups to disk. The non-traditional customer (right), meanwhile, also begins with 1 TB of data and adds 20 GB of new data each week. But this customer also leverages capacity reduction; data is both de-duplicated and compressed.

---

<sup>1</sup> ESG Research Report: *VTL Adoption and Market Trends*, March, 2007.

Figure Two: Power of Capacity Reduction

Week	Without Capacity Reduction				With Capacity Reduction			
	Primary data	New data	Backup data amount	Physical Storage Required	Primary data	New data*	Backup data amount**	Physical Storage Required (after de-dupe and 2:1 compression)
1	1,000	--	1,000	1,000	1,000	--	1,000	500
2	1,000	20	1,020	1,020	1,000	20	1,020	520
3	1,020	20	1,040	2,060	1,020	20	1,040	540
4	1,040	20	1,060	3,120	1,040	20	1,060	560
5	1,060	20	1,080	4,200	1,060	20	1,080	580
6	1,080	20	1,100	5,300	1,080	20	1,100	600
7	1,100	20	1,120	6,420	1,100	20	1,120	620
8	1,120	20	1,140	7,560	1,120	20	1,140	640
9	1,140	20	1,160	8,720	1,140	20	1,160	660
10	1,160	20	1,180	9,900	1,160	20	1,180	680
11	1,180	20	1,200	11,100	1,180	20	1,200	700
12	1,200	20	1,220	12,320	1,200	20	1,220	720
13	1,220	20	1,240	13,560	1,220	20	1,240	740
Total quarterly				<b>13,560</b>				<b>740</b>

\*\*Only non-duplicate data is backed up. Source: ESG

\*Assumes 2% change rate.

In this example, the traditional method would require about 13.5 TB of physical storage capacity over the 12-week period, compared to 0.740 TB in the capacity reduction environment. Naturally, the capacity reduction ratio will vary depending on the application, the amount of unique data and the number of full backups performed.

These same efficiencies can also be applied to the remote backup process because only new and unique data is moved over the WAN. ExaGrid claims up to a 50:1 efficiency in remote replication. This equates to 20 TB of data being replicated over a WAN while transferring only 400 GB of actual data.

Figure Three: ExaGrid Disk-Based Data Protection

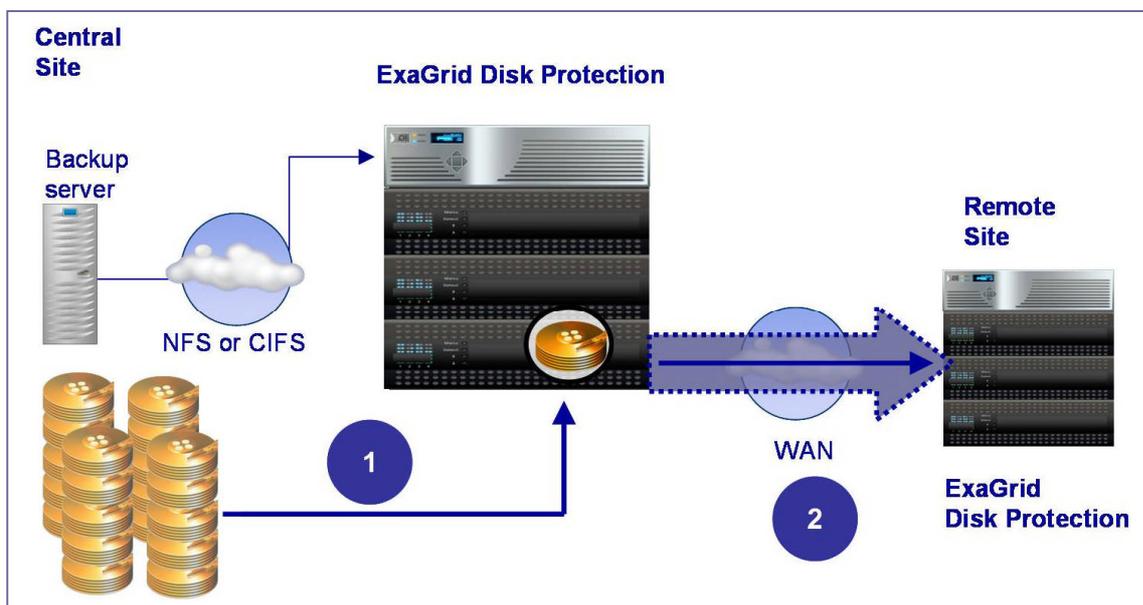


Figure Three illustrates the benefits of the ExaGrid solution. 1.) ESG has found customers often back up the same data over and over again when performing full backups. However, much of this data has already been backed up with the previous full backup done the week before. ExaGrid only stores the unique new data being backed up (again, data is de-duped at a byte-level) and therefore can greatly reduce the amount of capacity required. 2.) Not only is there a capacity benefit, but this optimization technology also reduces the bandwidth required to replicate this data over the WAN. Customers reduce their WAN costs and improve data protection.

## Other Considerations

### The Green Effect

Power, cooling and data center efficiency are definite “hot buttons” today. The proliferation of energy-inefficient systems has put a strain on power grids to the point that there are literally data centers with servers sitting idle because there is no available electricity to power them. ExaGrid’s capacity optimization technology allows for a significant reduction in the amount of backup data, hence, the number of disk drives to house the same “effective” capacity. Put simply: Less storage capacity means fewer disk drives to keep spinning and smaller systems to power and cool.

### Security

Similarly, capacity optimization also addresses the sensitive topic of information security. The less information that is out there, the less susceptible it is to attack and potential breach. If a customer list is e-mailed to 10 salespeople and backed up each time, it’s theoretically 10 times more vulnerable. Using capacity optimization technology, this risk drops considerably. In addition, ExaGrid actively employs a number of other secure methods throughout the process of backing up customers data, including the use of encrypted VPNs when transferring data between primary and off-site systems. And, of course, by backing up to disk versus tape, users no longer have to worry about the mishandling or loss of physical tape cartridges.

### Centralized Disaster Recovery

What disk-based backup brings to the table above and beyond using remote replication provided by a primary storage system is that it can protect ALL of your data. Data can then be replicated to a remote site for disaster recovery purposes. Applying ExaGrid’s capacity and WAN optimization techniques provides a single solution for all storage that is centralized and cost-effective. Imagine being able to back up 10 TB of data while only transferring 200 GB of data across the WAN to the remote site.

### Reliable and Recoverable

ExaGrid supports RAID 6, which protects against a dual-drive failure in a parity group. This is an extremely high level of protection, which is important as drives increase in capacity and therefore store more and more data. Additionally, ExaGrid conducts data integrity checks to make sure that files are stored and recoverable. ExaGrid performs this as a background task to ensure that when you have to restore data, it actually can be recovered. ExaGrid combines data integrity and recoverability at the drive level with RAID 6, and at the file level through proactive analysis.

### Issues and Challenges

ExaGrid has done a great job building momentum and raising awareness, but the job is not nearly done. The D2D backup market is heating up and ExaGrid needs to be on as many shortlists as possible. It has to continue to drive its momentum and expand its channel.

### The Bottom Line

The value that ExaGrid solutions provide transcends that of traditional tape-based backup solutions. ExaGrid offers superior backup and recovery times over tape as well as other less “intelligent” disk-based systems. ExaGrid was one of the first vendors to offer capacity optimization (in the form of byte-level de-duplication and compression) to address some of the biggest issues facing organizations today.

To recap, ExaGrid's disk-based backup system provides a number of clear benefits to users:

- Organizations can reduce capital expenditure on disk capacity by putting off disk purchases.
- Organizations can replicate data much faster and more efficiently to reduce overall WAN costs.
- Organizations can keep data on disk longer (i.e., they can lengthen retention periods).
- Organizations can lower data center environmental costs.

ExaGrid has also done a great job in the last year of not only vocalizing the benefits of its solution, but also educating users about disk-based backup in general and capacity optimization specifically. ESG attributes much of this success to a team that really lives and breathes the importance of these capabilities. Yes, success requires a good product, but also needs the ability to execute. The landscape is hot and busy and the competition fierce, yet ExaGrid is still in an excellent position to continue to gain traction in the burgeoning D2D backup market. ESG believes that ExaGrid is well-positioned to be a leader in this rapidly emerging market.