

The Economics of Backup

Why Moving to Disk-Based Backup Makes Good Business Sense

December 2006



For years, organizations have relied on tape as the primary method of backing up critical data. However, managing tape and restoring data from it can be extremely cumbersome, unreliable, time consuming and even costly. Since the cost of disk has fallen, many organizations are turning to disk for faster and more reliable backups and restores. Affordable disk-based backup products are arriving on the market at a time when data is rapidly and continuously growing, offering organizations of all sizes the opportunity to take advantage of a more cost-effective technology that helps to streamline IT operations.

Vital Company Data at Risk

Most organizations perform backups using a software product such as CommVault Galaxy, CA BrightStor ARCserve Backup, Symantec Backup Exec, NetBackup, or others which write to tape media as well as to disk. But, if a tape fails and data is lost, the data can be recovered only up to the point where the last backup occurred. Additionally, writing to sequential tape is slow, and the process taxes the application servers and network, which dramatically degrades user performance. To avoid performance degradation, backups are typically performed at night so as to not impact user productivity. And as data grows, so does the time it takes to back up, and the “nightly backup window” grows beyond the evening hours, endangering user productivity. If a backup is still running when users arrive at work in the morning, it must be aborted so that users can work with the application server and the network at full performance. If the backup is aborted, a complete copy of the data or all changed data does not exist, risking data loss for the previous 24 hours. If a file is deleted, data is corrupted or a drive fails, it’s possible that recovered data may be days or weeks old.

Examining the Nightly Backup Process

Tape libraries continue to write at faster and faster speeds in order to keep up with ever-growing amounts of data and fixed nightly backup windows. Additionally, tape libraries use exposed heads to write to tape as it streams by. But if the heads are not kept clean, the data will not write. Tape backups can present a number of problems, including:

- Blank Tapes– the backup ran but no data was written;
- Corrupted Files– where some of the data is recoverable and other data is not;
- Missing Files– when recoveries occur, some files exist but cannot be recovered; and
- Jammed Tapes – where backups don’t run, so an organization’s last good backup may be days old.

Since tape is a physical medium, there is no way to know if all data is present on the tape or if the data is recoverable until a recovery is needed.

Industry-Wide, Approximately 30 Percent of All Recoveries Fail

When a recovery is performed and the data is missing or corrupted, the IT administrator must roll back to older tape sets to find a tape set that is good. This takes an incredible amount of IT resource time. Additionally, as the IT administrator rolls back to older tape sets, users lose more and more data. If the last good tape run was three days ago, then users have lost three days of work. This work could be schematics, architectural drawings, CAD files, financial documents, legal documents, spreadsheets, and even worse, business emails or database entries with critical customer information.

Falling Disk Prices Make Disk-based Backup a Reality

Fortunately, disk prices have fallen in recent years, and it is now possible to use the same fast, reliable disk that has been used for primary storage for backups. Using disk for backups is now approximately the same price as tape, and the speed and reliability is far greater. Additionally, disk has numerous other advantages, including:

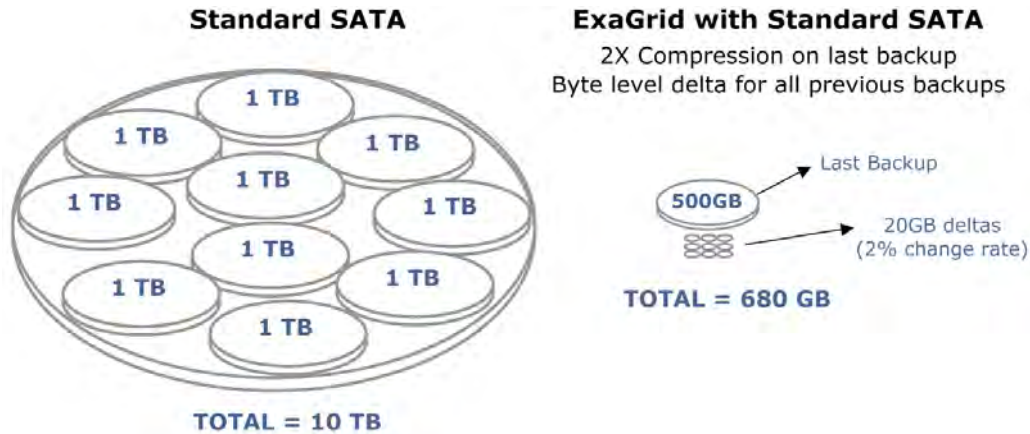
- Is a sealed media which requires no read/write head maintenance;
- Provides higher performance, reducing the nightly backup window;
- Provides greater reliability, virtually guaranteeing that data will be available for restores;
- Is far more secure than tape. Disk resides in a data center rack behind data center and network security; and
- Virtually eliminates manual intervention by IT staff, saving valuable IT resources and eliminating the possibility of human error.

Cost-Effective Disk-Based Backup

Several years ago, SATA (serial ATA) drives became highly available and highly reliable at about one-fourth the cost of standard primary storage disk. This one event began a series of technological changes that have driven the cost of disk-based backup down. A key factor in driving the cost of disk-based backup down is that on disk, data can be compressed and used alongside byte-level data de-duplication technology to significantly reduce the amount of data stored. When the data is both compressed and then reduced, the amount of data that can be stored on disk is far greater than tape. As a result, backups and restores are significantly faster, and the amount of disk space required is dramatically reduced.

Backing up data with tape can be extremely redundant. If your organization has 1TB of data to backup from week to week, the industry average is that about 2 percent or 20GB, of that data actually changes at the byte level. Some files are added, some are deleted, some are changed, and the majority has not changed at all between backups. In traditional backup scenarios, the backups are stored over and over, even when only 2 percent of the bytes have changed. Storing unchanged bytes is highly inefficient.

Example: 1TB of data, 10 full backups, 2% change rate



The Economics of Byte-Level Data De-Duplication

Backing up data to disk is far more efficient when byte-level data de-duplication technology is deployed, minimizing the amount of data to be stored. Byte-level data de-duplication stores only changes from backup to backup instead of storing full file copies. For example, if a full 1TB backup is performed over ten weeks, 10TB of data is backed up. But, with a disk-based backup product that uses both 2 to 1 standard data compression for the most recent backup and byte-level data de-duplication for all previous backups, the most recent 1TB backup is compressed 2 to 1, to 500GB, and only 20GB of byte level changes are stored for each previous backup, for about 180GB of byte level changes. The total requires 680GB of compressed and data-reduced data, or about 1/15th of the original data that needs to be stored, saving not only disk space, but reducing the cost of long-term data retention.

Improved IT Productivity and Reduced Risk of Failed Backups

Fortunately, all major backup applications natively support writing to disk in addition to tape. This means that an IT organization can add highly reliable disk instead of tape while continuing to use its existing backup application. Widespread adoption of using disk for backup has already begun. It is estimated that billions of dollars of disk have already been deployed in the backup process as disk-staging, and now with compression and data de-duplication techniques emerging, tape backup is quickly on its way to extinction. As companies seek to upgrade their tape libraries, they are realizing that for about the same amount of money they can move to disk and shorten their backup window, ensure reliable data recovery, free up valuable IT staff time and increase data backup security.

Organizations can use a disk-based backup system onsite and still use tape for offsite retention or disaster recovery. Disk can also be deployed across two sites, eliminating tape altogether. In fact, more than 70 percent of disk-based backup customers that use a system with compression and data de-duplication entirely eliminate tape with a two-site disk-based backup system.

About ExaGrid Systems, Inc.

ExaGrid is a leader in cost-effective disk-based backup solutions. ExaGrid understands the IT environment, and has created a complete, turnkey solution that is highly cost effective and fully supported.

The ExaGrid Solution

Is turn-key: all hardware and software is included, simply plug into the backup server and perform disk-based backups. Uses your existing backup application – no work, no change, no risk, no retraining

Is easy to use: all backups and restores are still controlled by your existing backup server

Is cost effective: with two to one compression and byte level data de-duplication the cost is equivalent to backing up to tape

Provides a two-site solution: moves only the bytes that change between backups over the WAN ensures offsite data on disk for data redundancy and for quick recovery from site disasters

Incorporates brand components: motherboard built by Intel with Intel Xeon processors, SATA drives built by Seagate Technology, the leading disk drive manufacturer

Is fault resilient: RAID6 with a spare drive can survive two simultaneous disk-drive failures

Has hot swappable drives: Dual hot swappable power supplies

Includes 100% support and maintenance: email and phone support, all point and full version releases (all new features). Next business day replacement of any failed hardware components. ExaGrid's unique monitoring of alerts and alarms is included

About ExaGrid

ExaGrid® offers a turnkey appliance that works in conjunction with your existing backup applications and is 30 percent the cost of standard SATA disk. ExaGrid provides next generation byte-level data de-duplication, which stores only byte-level changes for each version instead of storing full file copies. This unique approach reduces the amount of disk space needed to at least 20 to 1, resulting in a significant cost savings over standard SATA storage.

Compliments of: 

 **ESS** ENTERPRISE
Storage Solutions

3835R E. Thousand Oaks Blvd. #315
Westlake Village, CA 91362
Tel 877.230.2837 ~ Fax 805.435.2500 ~ www.ESS-Direct.com