APPLICATION NOTE

Using BladeStore disk systems with Tivoli Storage Manager software

A valuation method for determining when disk-to-disk backup and restore make sense in your environment

AUGUST 2004

1 Executive summary

BladeStore disk systems from StorageTek® are based on serial advanced technology attachment (SATA) and designed to satisfy real-time data needs with maximum reliability and availability. Tivoli Storage Manager (TSM) software, an IBM product, provides data protection to a wide variety of open systems operating environments. You can combine these components to create a disk-to-disk backup strategy that can complement your existing tape backup and restore systems. Disk-to-disk backup offers higher performance than traditional tape systems. For example, disk systems can eliminate the wait times associated with tape cartridge mounting and positioning. You also have the flexibility to perform multiple, simultaneous backup and restore operations because of the inherent capability of disk systems to read and write data simultaneously. A disk-to-disk approach increases reliability, data availability and improves operational efficiency compared to tape systems.

A BladeStore/Tivoli disk-to-disk backup strategy is considerably less expensive than disk mirroring. Mirroring may be justified in environments requiring near continuous data availability. However, because a disk mirroring approach requires the duplication of disk devices, the cost may not be justified for all types of data. A disk-to-disk backup strategy is a lower cost alternative to primary disk storage. It also provides faster data recovery than a traditional tape backup solution and can be implemented for a broader range of data.

The strongest case for a disk-to-disk backup or restore strategy occurs when the time it takes to perform backup or restore exceeds their expected windows, causing applications to be unavailable and resulting in a loss of productivity. Backup or restore overruns may occur for several reasons, including increased LAN traffic, inefficient device drivers or the wait times associated with tape mounting, positioning and data transfer. If these delays result in a critical application being unavailable on a regular basis, you can determine the number of hours and the potential revenue lost over the course of a year. The value of reclaiming a portion of this revenue by reducing the backup or restore window can more than justify the investment in the BladeStore disk system.

In addition, restoring databases and file systems is much quicker from disk. You can store multiple days of backups on a single disk, and, because most restore operations involve the most recent daily backups, you can expect shorter business recovery times. This is particularly true for data that would span multiple tape volumes. That's because a disk-to-disk strategy eliminates the delays associated with tape mounts, positioning, rewinds and dismounts.

2 Shrinking backup windows

Storage managers have seen a consistent trend of rapidly growing data volumes combined with the need for 24 x 7 x 365 data accessibility. The result is the ongoing challenge of how to back up and restore increasingly large amounts of data within shorter and shorter time slots. Tape systems have been the backup and restore option of choice, because they offer a cost-effective way to handle large volumes of data at reasonable speed. However, shrinking backup windows are straining the performance limits of tape systems to keep pace.

For critical data with the highest availability requirements, storage managers have justified the use of expensive primary disk as the backup media. Recently, lower cost disk technology such as SATA has made disk systems an affordable alternative, even for data that does not require continuous or near-continuous availability.

If reducing the backup window is your primary concern, StorageTek's BladeStore disk systems, combined with the Tivoli Storage Manager software, may be able to speed up your backup and restore process compared to the use of tape systems, and at a lower cost than primary disk systems used for the same purpose. You can insert the BladeStore disk systems as an intermediate or "secondary" layer in the backup hierarchy between primary disk storage devices and the secondary tape storage device. This enables backups and restores to occur directly to or from the secondary BladeStore disk system.

3 BladeStore/TSM disk-to-disk solution overview

IBM's Tivoli Storage Manager (TSM) software provides comprehensive, integrated data protection, business continuance and archive management, through a scalable client-server architecture capable of backing up and restoring hundreds of enterprise-class servers.

3.1 TSM software architecture

The TSM software architecture differentiates it from other network backup and restore solutions. Unlike data protection software that uses a flat-file index or catalog to track the location of backups, TSM uses a relational database and recovery log. The flat-file index or catalog approach restricts the ability to move backup objects from one type of storage device to another. This makes it more difficult for you to manage your backed up data effectively. In addition, with the current rapid growth in data volumes, a flat-file approach carries the risk of exceeding the capacity of the file system in which it resides. If that happens, you would need to move the flat-file index or catalog to a new file system with a larger capacity.

Compared to these limitations, TSM's underlying relational database is much more scalable and adaptable to changes in your data environment. The TSM database and log can span multiple volumes in different file systems, adding to the scalability of the TSM solution. Storage administrators also have the flexibility to define storage management policies for individual files, clients or group of clients. They can track the number of versions, storage destination, retention periods and location on storage media on an individual file level.

The TSM solution uses two-phase commit processing which stops partial file references from being created as a result of a failed backup/archive transaction. Instead, TSM provides the ability to mirror the database and recovery log. The mirrors can be taken online or offline and resynchronized as necessary without affecting the availability of the TSM server as a protection against hardware failures.

3.2 Backup methodology

With traditional backup approaches, a full backup is performed on a weekly basis with daily incremental backups that replicate only the files that have changed since the last full backup. This can result in redundant full backups of files that have not changed over the course of several full backups. Restoring data typically involves restoring the most recent full backup and then overlaying the most recent incremental backups by copying multiple files in multiple steps. This process is time consuming and inefficient for both backups and restores.

In contrast, the TSM software uses a "progressive incremental" approach that starts with a full backup, but then involves backing up only the new or changed files in each subsequent backup. The TSM relational database updates the current backup image of any file that has changed. That means only a single complete copy of the backed up data exists at any time. Restoring the data simply involves copying the single backup image to the client. If a restore process fails, TSM can resume the process from the point of failure.

The TSM approach eliminates weekly full backups and redundant copies of the same files. It also reduces the backup and restore times, reduces the corresponding data traffic on the network and backup servers and reduces storage device and media requirements.

3.3 Media usage

With the TSM approach, you can move data from one storage device to another. This sets TSM apart from traditional data protection architectures. You can backup data to disk and then selectively migrate some or all of the data to a single tape or set of tapes. TSM uses an efficient approach for logging the location of data on tape, avoiding potential delays during restore operations. TSM also provides a tape reclamation capability to consolidate data stored on tape and reclaim space occupied by expired files.

3.4 Supported operating environments

Tivoli Storage Manager with BladeStore is compatible with a variety of host operating systems as shown in the following table:

Environments supported by TSM

TSM server	TSM client
HP®-UX 11.0 or 11.11	HP-UX 11.0 or 11.i (32/64 bit)
IBM® AIX 5L 5.1 or 5.2	IBM AIX 5.1 or 5.2 (32/64 bit)
Red Hat Enterprise Linus AS 2.1, ES 2.1, WS 2.1	Linux x86 2.4 kernel (Red Hat 7.2, 7.3, 8.0, Advanced Server 2.1; SuSE 7.3, 8.0, 8.1, SLES 7 and 8; Turbo Linux 7.5, 8.0)
Windows® 2000 Professional, Server, Advanced Server, Datacenter Server	Linux pSeries 2.4 kernel (SuSE 8.0)
Windows Server 2003- Standard Edition (32 bit), Enterprise Edition (32/64 bit), Datacenter Edition (32/64 bit)	Linux/390 and zSeries 2.4 kernel (SuSE Enterprise Server 7, 8)
Sun® Solaris 8 (64 bit) or 9 (64 bit)	SGI® IRIX 6.5 with EFS or XFS Sun Solaris 7, 8, 9 (32/64 bit)
	Windows XP (32/64 bit), Windows Server 2003 (32/64 bit), Windows 2000 Professional, Advanced and Datacenter Server
	Windows NT 4.0 SP5, SP6a
	Novell® NetWare 5.1, 6

Table 1. Tivoli Storage Manager supported environments

4 When does disk-to-disk backup make sense?

How do you know when the costs of implementing a disk-to-disk backup strategy are justified? You can quantify the value of a BladeStore disk-to-disk backup strategy through its benefits in reducing the time required for backups and restores. You will need to characterize the current performance and the costs incurred from elongated windows for both backup and restore operations.

4.1 Quantifying the potential value of reducing the backup window

A number of factors contribute to elongated backup windows, including:

- Wait times for tape mounting, positioning and data transfers
- Heavy LAN traffic
- Lack of sufficient server memory
- Outdated device drivers
- Inefficiencies in backup policies and/or in backup software.

Admittedly, LAN traffic, server memory and device driver issues can affect the backup window regardless of the technology used. Assuming those factors are constant, the justification for using a disk-to-disk strategy to reduce backup windows can be built with the following steps:

1. Identify the service levels for backups and the variances to those service levels.

   This assumes you have formal service level agreements defining backup performance. If no such agreements exist, you can define the backup expectations and current status by answering these questions.

   - What applications are being backed up?
   - What is the frequency of backups?
   - How long is the backup window?
   - How frequently do backups exceed the backup window?
   - By how much do the backups exceed the backup window?

2. Characterize the backup environment.

   This information will help identify how BladeStore can best be used to reduce the backup window. Asking the following questions will help you characterize the backup environment.

   - How much data is being backed up?
   - How many tape drives are being used?
   - What types of tape drives are being used?
   - Are backups being multiplexed?
   - How many simultaneous backup streams are running at any one time?

3. Determine the employee productivity and financial impact of elongated backup windows.

   If employees can't access application data because the data is being backed up, their productivity goes down and company revenue suffers. Answering the following questions can help you characterize the impact of long backup windows on productivity and revenue.

   - How many people have access to the application(s) being backed up?
   - How often and how long do employees access the application(s)?
   - How much revenue is generated per person in the company?
   - How much revenue is generated per hour?

We can create a sample justification that answers the above questions for a hypothetical medical supply company. The company has 200 employees and annual revenue of approximately $350 million. The sales organization backs up its order entry application on a daily basis. The order entry application has grown to 16 gigabytes in size to support the company's expanding business. The company uses a single T9840A tape drive from StorageTek as the destination device for the backups. Because of the growing data volume, daily backups take an average of 35 minutes to complete, a 15-minute overrun compared to the allotted 20-minute backup window. The result is the order entry application is unavailable for 15 minutes every day, leaving 25 sales associates unable to process their sales orders.

The following table contains the key data for computing the cost of backup delays in this example.

Service level agreement and variances	
Application	Order entry
Backup window (hrs)	0:20
Backup frequency and type	Daily-full
Backup elapsed time (hrs)	0:35
Frequency of overrun	7
Backup strategy and configuration	
Amount of data backed up	16 GB
Number of tape drives	1
Type of tape drives	T9840A
No. of simultaneous backup streams	1
Employee productivity and financial impact	
Total employees	200
No. of employees accessing application	25
Annual revenue	$350,000,000
Revenue per hour	$80,128
Annual revenue per employee	$1,750,000

Table 2. Hypothetical values for calculating the cost of backup delays

In this example, the company backs up 16 gigabytes of data to a single T9840A tape drive with one 16-gigabyte I/O stream. The elapsed time to complete the backup is 35 minutes. In actual testing with Tivoli Storage Manager, a three-blade RAID5 BladeStore was able to backup 16 gigabytes of data with an approximate elapsed time of seven minutes. For the purposes of this example, we will assume that a BladeStore configured similar to the one in the TSM test environment will adequately handle the 16-gigabyte backup requirement.

Table 3 compares the backup performance in the company example with BladeStore and TSM.

Service level analysis		
	T9840A	**BladeStore**
Backup window (hrs)	0:20	0:20
Backup elapsed time (hrs)	0:35	0:07
Backup window overrun (hrs)	0:15	0:00
Backup window reduction (hrs)	0:00	0:27
Backup strategy analysis		
	T9840A	**BladeStore**
No. of backup I/O streams	1	1
Data transfer per I/O stream	16 GB	16 GB
Aggregate data transfer	16 GB	16 GB
No. of units	1	1
Productivity/Financial impact analysis		
No. of application users	25	
Annual revenue	$350,000,000	
Revenue per hour	$57,234	
Revenue per employee per hour	$191	
Application user hourly rate	$30	
	T9840A	**BladeStore**
Annual production hours lost due to backup overrun	(2,275)	0
Annual production hours gained due to backup reduction	0	1,972
Annual productivity loss due to backup overrun	($68,250)	$0
Annual revenue loss due to backup overrun	($435,028)	$0
Annual revenue gain due to backup window reduction	$0	$376,157

Table 3. Financial impact of reducing the backup window using BladeStore

The data in this example shows that the backup window was reduced by using the BladeStore disk-to-disk backup strategy — the elapsed time went from 35 minutes down to 7 minutes; a savings of 28 minutes or a 77 percent reduction in the elapsed time for the backup. The annual cost of the backup overrun is $68,250 in lost employee productivity and approximately $434,000 in lost revenue. By reducing the backup elapsed time, the backup window for the order entry application can be reduced by 13 minutes or approximately 65 percent. This is calculated based on the difference in backup window time and the reduction in elapsed time for the backup. As a result of the increased availability of the order entry application, annual revenue can be expected to grow by approximately $376 million. You can calculate the annual revenue gain by multiplying the annual production hours gained (for the 50 CRM users) due to backup window reduction by the average revenue per employee per hour.

4.2 Quantifying the potential value of reducing restore times

In addition to backing up critical data in a timely fashion, storage administrators must have processes for restoring the data quickly whenever necessary. In the example of the small medical supply company, the order entry application experiences unexpected outages due to data corruption approximately twice a month. When the corruption is discovered, the IT staff must restore the order entry application to the most recent good backup. With some companies this may require restoring the last full backup along with restoring additional incremental backup files to arrive at a completely restored backup copy. Such restore processes can be lengthy. In contrast, the BladeStore/TSM process requires only a simple restore of the last full backup copy, since a full backup is done on a daily basis. In either case, of course, additional work will be required to re-input all the lost sales orders from the time of the last backup to the time when the data corruption occurred.

In this example, a service level agreement for the order entry application specifies that a restore should take no longer than 30 minutes to complete. However, because of the growth of the order entry application, restores are now requiring approximately 40 minutes. And during that time the 25 sales associates are unable to process sales orders.

Factors that affect elongated restores times are similar to those affecting backups:

- Wait times for tape mounting, positioning and data transfers
- Heavy LAN traffic
- Lack of sufficient server memory
- Outdated device drivers
- Inefficient backup policies and/or backup software.

Again, LAN traffic, server memory and device driver issues are assumed to be constant factors affecting restore times for both tape and disk-to-disk approaches.

In addition, restore operations for non-sequential data can take significantly more time, especially when the file(s) resides on different parts of the tape media or in some cases on a different tape media altogether.

You can estimate the value of reducing restore times with a BladeStore disk-to-disk strategy by using steps similar to those used to evaluate shorter backup windows, including:

1. Identify the service levels for restores and the variances to those service levels.

   Answering the following questions can help you characterize the requirements and expectations for restores and how they are being met currently.

   - What is the frequency and quantity of restores?
   - What are the service level criteria for restores of a single file, directory or volume?
   - How long do restores typically take?
   - What is the typical unit of storage being restored (for example, a single file, directory or volume)?
   - How frequently do restores exceed the service level criteria for restores (or the expectation of users)?
   - By how much do restores typically exceed the service level criteria or user expectations?

2. Characterize the restore environment.

   Answering the following questions can help identify how a disk-to-disk strategy can be implemented to reduce restore times.

   - What is the typical restore (single file, directory or volume)?
   - How much data is being restored?
   - How many tape drives are being used?
   - What types of tape drives are being used?

3. Determine the employee productivity and financial impact of elongated restore times.

   Answering the following questions will help identify the impact of current elongated restore times on application availability and in turn on productivity and revenue.

   - How many people have access to the application(s) being restored?
   - How often and for how long do employees access the application(s)?
   - How much revenue is generated per person in the company?
   - How much revenue is generated per hour?

With the answers to these questions, you can build a justification for a BladeStore/Tivoli Storage Manager disk-to-disk strategy for reducing restore times similar to the one in Table 4. You would be restoring data directly from disk and thereby allowing single file, directory or full volume restores to complete in a much shorter time. End user productivity could be improved by reducing the amount of time that the user must wait while data is being restored.

Service level agreement and variances	
Application	Order entry
Backup frequency and type	Daily-full
Backup frequency	Twice monthly
Restore service level (hrs)	0:30
Restore elapsed time (hrs)	0:40
Frequency of overrun	2
Backup strategy and configuration	
Amount of data backed up	16 GB
Number of tape drives	1
Type of tape drives	T9840A
No. of simultaneous backup streams	1
Employee productivity and financial impact	
Total employees	200
No. of employees accessing application	25
Annual revenue	$350,000,000
Revenue per hour	$80,128
Annual revenue per employee	$1,750,000

Table 4. Hypothetical values for calculating the cost of restore delays

In this example, the company restores 16 gigabytes of data from one T9840A tape drive with a single 16-gigabyte I/O stream. The elapsed time to complete the restore is 40 minutes. In actual testing with Tivoli Storage Manager, a three-blade RAID5 BladeStore was able to restore 16 gigabytes of data with an approximate elapsed time of 7 minutes. For the purposes of this example, we will assume that a BladeStore configured similar to the BladeStore in the TSM test environment will adequately handle the 16-gigabyte restore requirement. Table 5 compares the customer's restore environment with BladeStore.

Service level analysis		
	T9840A	**BladeStore**
Restore window service level (hrs)	0:30	0:30
Restore elapsed time (hrs)	0:40	0:07
Restore window overrun (hrs)	0:10	0:00
Restore window reduction (hrs)	0:00	0:32
Backup strategy analysis		
	T9840A	**BladeStore**
No. of restore I/O streams	1	1
Data transfer per I/O stream	16 GB	16 GB
Aggregate data transfer	16 GB	16 GB
No. of units	1	1
Productivity/Financial impact analysis		
No. of application users	25	
Annual revenue	$350,000,000	
Revenue per hour	$57,234	
Revenue per employee per hour	$191	
Application user hourly rate	$30	
	T9840A	**BladeStore**
Annual production hours lost due to restore operations	(400)	(70)
Annual production hours gained due to restore time reduction	0	330
Annual productivity loss due to restore operations	($12,000)	($2,100)
Annual revenue loss due to restore operations	($76,313)	($13,355)
Annual revenue gain due to restore time reduction	$0	$62,958

Table 5. Reducing restore times with BladeStore

Most companies will strive to restore data as quickly as possible, so their primary measure is the actual elapsed time for a restore operation. In this example, the restore elapsed time was reduced from 40 minutes to 7 minutes, a savings of 33 minutes or 83% in elapsed restore times. The tape-based restore operation costs the company approximately $12,000 annually in lost employee productivity. In contrast, the BladeStore solution reduces the employee productivity loss by 82.5% or $9,900 on an annual basis while increasing potential revenues by approximately $63,000.

The example of our hypothetical medical supply company makes several assumptions about order entry application use and its impact on company revenue. In your own company, you may have slightly different assumptions, even if you are using an order entry application to justify a disk-to-disk strategy. But the justification process outlined previously would be the same regardless of the number of users impacted by backup or restore window overruns, or the type of critical, revenue-generating application.

5 Measuring actual benefits of disk-to-disk backup and restore

Once you've implemented a disk-to-disk strategy, it is critical to measure actual operational improvements. This will help you learn from the implementation and provide valuable metrics for future disk-to-disk backup and restore investment decisions. Every backup environment is unique — the best rules of thumb are those generated from actual experience.

The BladeStore disk-to-disk strategy offers a low-cost disk solution as a secondary storage device within your backup process that enables you to meet your service level agreements for backups and restores. In addition, BladeStore reduces the potential impact of tape-related backup failures and can help minimize the burden on your administrators.

In the medical supply company example described above, the BladeStore approach provided the following benefits:

- A 77 percent reduction in backup elapsed time
- A 65 percent reduction in the application backup window
- Improved employee productivity resulting in potential revenue gains
- Fewer backup failures and lower risk
- Decreased burden and risk for system administrators

Similar additional benefits result from using the BladeStore disk-to-disk strategy for restore operations, including:

- Reducing restore times by approximately 83 percent
- Eliminating tape delays (mounting, positioning, etc.)
- Increasing employee productivity and minimizing lost revenue due to data unavailability
- Decreased burden and risk for system administrators

Your StorageTek representative can help you define the necessary configuration for the amount of data you are backing up and the expected performance for comparison with your current backup and restore strategy.

Backup and restore operations will continue to be an issue for IT managers. A BladeStore/Tivoli Storage Manager system is a viable solution for solving the problem of shrinking backup/restore windows. IT offers you the opportunity to optimize your backup processes, reduce data recovery times, increase the reliability and availability of critical information and reduce the burden on your operations staff. From a business perspective, reduced backup and restore times equate to increased employee productivity and revenue.

ABOUT STORAGETEK®

Storage Technology Corporation (NYSE: STK) is a $2 billion global company that enables businesses, through its information lifecycle management strategy, to align the cost of storage with the value of information. The company's innovative storage solutions manage the complexity and growth of information, lower costs, improve efficiency and protect investments. For more information, visit www.storagetek.com, or call 1.800.275.4785 or 01.303.673.2800.

WORLD HEADQUARTERS

Storage Technology Corporation
One StorageTek Drive
Louisville, Colorado 80028 USA
1.800.877.9220 or 01.303.673.5151

DT 0001 A 08/04