WHITE PAPER

Optimizing data protection

Applying information lifecycle management to raise protection service levels with fewer resources

Conventional data protection methodologies are insufficient for today's high-velocity business environment. Backup and recovery processes are slow, unreliable and require too much staff time. A total systems approach is needed to align disparate protection processes and realize economies of scale. By employing information lifecycle management, data can be protected according to its purpose and value, improving data protection efficiency and simplifying backup and recovery processes.

APRIL 2004

StorageTek®

1 Executive summary

> Faced with flat budgets, IT organizations have to find innovative ways to address rapidly escalating data protection challenges.

> Information lifecycle management solves data protection challenges by aligning storage options with the purpose and value of data.

> A total systems approach to data protection can be implemented today following a three-step process: assess assets, adapt the infrastructure, monitor and tune protection systems.

> Information lifecycle management principles have helped organizations achieve faster and more reliable backup and recovery with reduced labor costs.

2 The challenge

2.1 Optimizing data protection — without starting over

For most IT departments, data protection is a daily struggle. People-intensive, manual processes can overwhelm staff. Backup and restore processes are often unreliable. Current systems can't adequately protect growing volumes of data. Backup is made more complex by multiple operating platforms, inconsistent management tools and changing technologies.

These types of data protection challenges are widespread. In the words of the META Group, "Backup/recovery remains a primary pain point for most storage support organizations." The research firm estimates that 60 percent to 70 percent of the effort associated with storage management is related to backup/recovery.[1]

Current data protection challenges will only be exacerbated by continuing double-digit growth rates for data. In 2003 alone, data volumes grew at an estimated rate of 40 percent to 75 percent.[2]

For many enterprises facing this dramatic growth in data, conventional data protection tactics are providing only partial solutions to much larger problems. Yet faced with flat or even declining budgets, most IT organizations can't respond by throwing money at the problem. Even if they could, it would only increase the complexity of the storage protection infrastructure — worsening a root cause of the challenges they face. Instead, they need a fresh approach to data protection, an approach that unleashes the value of current storage infrastructures and makes maximum use of lower-cost storage options.

Information lifecycle management provides this new approach. It offers a strategy for enhancing the efficiency and cost-effectiveness of data protection. This strategy focuses on aligning identified data protection needs with optimal storage solutions, and then making maximum use of automation and standard processes to maintain appropriate protection levels as data values change.

3 Information lifecycle management

3.1 A total systems approach

A total systems approach to data protection looks beyond individual backup systems or data copies. It treats data protection as a whole, rather than as a collection of discrete systems. It recognizes that backup systems are interdependent, and that data should move easily from one system to another.

This holistic approach gets to the heart of the well-being of a business. It encompasses backup and restore, disaster recovery and business continuity processes, along with data center operations management. It focuses on improving performance, enhancing protection and generating economies of scale.

With a total systems approach to data protection, not all information is treated equally — a common shortcoming of today's backup processes. Instead, information is protected based on its purpose and value. The use of a tiered protection methodology strives to lower the total cost per terabyte of protected data.

This tiered methodology matches data protection needs with the value and purpose of data. Information with higher value and higher availability requirements is protected with higher performance storage and management options. Information that is less essential or less urgent to the business is protected with more affordable, lower performance storage options.

[1] "Magnetic Tape: Whither Thou Goest?" *META Delta*, September 15, 2003.
[2] Giga Research (part of Forrester Research), 2003

A tiered approach to data protection might take the form of Figure 1.

Under this approach, data remains protected and available, despite continuing growth and changing business requirements. But it is protected in a more strategic way, rather than through blanket backups of information regardless of its value.

A total systems approach to data protection consolidates administration and speeds up problem resolution. Backup and recovery processes are managed at a higher level. Management tasks are simplified by automating and integrating replication methods. Backup and recovery times can be reduced or eliminated for critical data without overwhelming budgets.

For lasting improvement to data protection in complex IT environments, it's not enough to add or update individual components or to improve existing management procedures. Information lifecycle management provides a total systems approach that aligns the pieces so they work together more effectively.

4 Applying information lifecycle management to data protection

4.1 A three-step process

Information lifecycle management can be applied today to enhance the efficiency of data protection solutions. This approach does not necessarily require new hardware, software or technology. It can be implemented by making strategic improvements to an existing data protection infrastructure.

An implementation of this type encompasses three steps:

- Assess information assets and uses.
- Adapt the data protection infrastructure to meet business needs.
- Monitor and tune data protection processes over time.

Each phase in the process includes specific steps that support broader business goals.

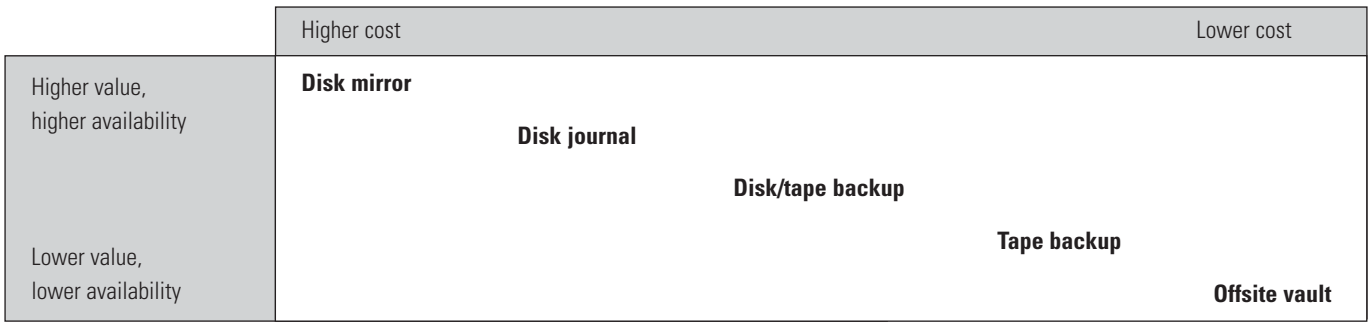

Figure 1. Matching data value with protection options.

4.2 Step one: assess information assets and uses

4.2.1 Classify information

A good first step in the assessment process is to categorize information according to its purpose and relative value. The classifications will span from data that is essential to the mission of the business to data that is merely convenient to retain.

Figure 2 shows an example of a classification system that groups data into one of four designations — critical, vital, sensitive or required. Some organizations classify their information as Gold, Silver and Bronze; others have 20 classification levels. Your system should be based on how your business uses its information.

Critical	Data that is critical to the company's ability to function Data that is used in key business processes
Vital	Data a company can function without for short periods of time Data used in standard business processes Data that represents a significant investment and that could be difficult to regain
Sensitive	Data used in everyday operations, but there are alternate sources in case of loss Data that can be reconstructed with relative ease
Required	Data that can be reconstructed with minimal cost Data that is already copied

Figure 2.

After data has been grouped into categories based on relative worth to the organization, go back to each group of data and document the risks the data is exposed to on a day-to-day basis.

In particular, look for two types of exposure:

- Probable risks, most of which can be found by looking back over events in the recent past
- Possible risks, or those that are unlikely but could cause the loss of data should they arise

At this point, one shouldn't assume any level of data protection. The goal is to identify all of the threats to different classes of data. The insights gained in this exercise will be valuable as the assessment process continues.

4.2.2 Evaluate current approaches

Storage is purchased primarily for its ability to deliver data where and when it is needed, whether it is being written into storage or read from it. To validate current data protection approaches, one should evaluate their ability to classify and retain that data. Where are various data objects housed? Is all of the most valuable data stored in enterprise-class storage devices? Is the least valuable data overprotected, and therefore using up more than its share of budget?

Starting with the most critical data, follow information through its useful life to determine if it is being managed in a way that protects it from loss. Is it findable, movable and recoverable at each phase of its life? Does the speed at which it can be found, moved or recovered match business needs? Requirements vary throughout the information lifecycle. A new financial transaction and a three-year-old financial transaction are both important, but they have very different access requirements.

As data's age and use characteristics change, protection requirements change. So should the media upon which the data is stored. Migration between storage classes is becoming faster and easier. New storage management capabilities make it possible to dynamically blend storage (disk and tape, for example) for a particular data application. These combinations can leverage the best characteristics of each media to create efficient, high performing solutions matched to specific requirements.

Identify requirements and gaps

Storage is usually managed as part of a data processing solution, even though some components may be shared across environments. Assessing storage as a single protective asset is useful in finding data protection gaps. Look closely at any storage resource being shared by data applications with widely different backup/ recovery requirements. Verify that critical data is not under-protected and that non-critical data is not overprotected.

This holistic view exposes opportunities for greater efficiency. It may uncover excess capacity that could be used by another application with similar protection requirements. Multiple sets of data with similar

requirements may be able to be clustered together and protected with a single backup/restore process, saving storage cost and administration.

This view can be consolidated into a table format, as shown below in Figure 3. This can help isolate the protection requirements for different classes of data. How much time can be allowed between backups? How many seconds, minutes or hours worth of data could be lost without affecting the business? What is the minimum restore time for a certain application? Is a certain type of data important enough to ship off to a remote recovery site? The answers to questions such as these help identify the optimal data protection technology for different data — in terms of both performance and cost.

In addition, a table such as this can expose gaps and inconsistencies in data protection. One should look more closely at areas that have generated question marks concerning data protection, and then, where warranted, put corrective plans in place.

There might be empty boxes in the table. That doesn't necessarily indicate a data protection gap. Information that is of low value or easily replaced might not warrant data protection and the associated costs. But policies should be in place to delineate the types of information that will not be protected. This helps avoid the expenditure of IT resources on unnecessary data protection measures.

4.3 Step two: adapt the data protection infrastructure

The assessment step should yield a clear view of the data protection levels that are necessary for different types of data. Step two of the information lifecycle management process for data protection acts on this information. In this step, data protection infrastructure is adapted to protect information according to its purpose and changing value.

Protection policies may specify very high levels of data protection. Do you need duplicate copies of backup data? Is a remote copy necessary for disaster recovery? Are multiple point-in-time copies of the data required? Policy-driven requirements must be included in decisions you make about your overall data protection infrastructure.

The different levels of data protection to be considered here include:

- Automated tape backup for a baseline level of data protection and availability
- Disk-enhanced tape backup for environments with lots of data to back up within small windows
- Disk journaling for continuous protection by capturing every data change, enabling fast recovery to any point and eliminating backup windows
- Disk mirroring with storage failover for high data availability and elimination of backup windows

	Recovery point (how out of date data can become)	**Recovery time** (how long the data can be unavailable)	**Backup window** (allowable downtime for backups)	**Remote recovery** (disaster recovery requirements)
Data class 1 (Critical)				
Data class 2 (Critical)				
Data class 3 (Vital)				
Data class 4 (Vital)				
Data class 5 (Sensitive)				
Data class 6 (Required)				

Figure 3.

4.3.1 Tune backup and restore systems

The first action in Step two involves the tuning of backup and restore systems to meet varying data protection needs. This can include enhancing performance for more valuable data with disk-to-disk-to-tape backups, consolidating backups with similar requirements, and setting clear policies on data that will not be protected.

Automated tape backup/restore: This is the minimum requirement for most enterprise data. Automated tape backup/restore offers the management advantages of automation together with the cost advantages of tape storage. For enhanced availability in more complex server environments, automated backup can be coupled with tape drive sharing.

Disk-enhanced tape backup/restore: For higher-demand data, an automated tape solution can be enhanced with a disk component between the tape and the application to enable faster recovery times and the elimination of backup windows. The disk either acts like very fast virtual tape, or stages a secondary copy from which backups can be performed. Both methods have their place. Either way, disk-enhanced backup is the higher performance and higher cost option to basic automated tape backup.

It should be noted that adding disk won't automatically improve the speed of backup. Backup speeds can be limited by the I/O performance obtained from production systems.

Tuning backup/restore systems is not about throwing money at problems. The assessment process might suggest gains that could be made by consolidating backups into libraries or fewer libraries. The economies of a centrally managed shared automated tape system can improve backup and restore speeds and reliability without increasing costs. Similarly, the process might identify certain data that does not need to be backed up because it is relatively unimportant or easily replaced.

4.3.2 Tune data replication systems

In addition to addressing backup and restore systems, data replication systems should be tuned to meet varying information availability needs. This can include applying journaling technology for cost-efficient instant recovery, reserving the use of data mirrors for critical online information, and consolidating the data replication infrastructure.

Disk journaling: Disk journaling provides continuous protection by capturing every incremental data change, enabling fast recovery and eliminating backup windows. This approach applies a data management technique based on discrete event journaling. It uses a disk component to continuously protect data at the point of use for day-to-day user recovery needs and integrates with tape backup systems for long-term archiving and disaster recovery.

Disk mirroring: Disk mirroring promises instantaneous, near real-time failover for data required by mission-critical applications. If primary arrays become damaged, traffic is switched to secondary or mirrored backup arrays. It can also eliminate backup windows and the associated system downtime.

In considering data protection options, cost is a critical factor. When it comes to data replication solutions, one size does not fit all. The decision to mirror requires a careful review of the criticality of the application that is being protected. Application criticality will determine whether an organization can wait for tape restore or must adopt a disk mirroring solution.

Design for the future

Data is growing rapidly, and storage technologies are advancing quickly. This makes it imperative to design the storage infrastructure for the future, as well as for present needs. The infrastructure needs to scale and adapt as data protection needs change. One caveat for the design process: Don't confuse data protection with data archiving. Data retention and recall requirements may go beyond basic protection needs satisfied by backup or disaster recovery. Make sure both needs are covered.

4.4 Step three: monitor and tune data protection processes

Maintaining a total systems approach to information lifecycle management is an ongoing process. It's important to monitor data protection processes holistically and in real time. This helps achieve improved reliability and faster problem resolution.

Monitoring and tuning work should include backup applications and storage networking components, as well as storage devices. The human factor is also important. Are backup processes failing because of human error? Are data protection tasks burdening IT personnel? Could some of these tasks be automated?

In this ongoing phase in the ILM process, look for opportunities to automate and consolidate backup and restore management with storage resource management tools and backup resource management tools. Software and services are now available to centrally monitor and manage the complete data protection infrastructure, locally or remotely.

Throughout this process, keep in mind that management is only as good as the quality of the inputs. Are you measuring the right things? The process of asking and answering such questions will point to opportunities to further enhance the data protection infrastructure to improve efficiency and pull more value from storage investments.

5 The opportunity

The application of information lifecycle management principles has helped organizations around the world achieve notable gains in data protection. A few examples of these gains follow.

5.1 Shorter backup windows

Following information lifecycle management principles, a rental car company reduced its backup runtimes by 30 percent and increased system availability and performance for thousands of associates. These gains were accomplished in part with the implementation of virtual tape technology and software that enables multiple backup jobs to access a single tape.

5.2 Faster recovery

A healthcare insurance provider in a hurricane-prone region employed information lifecycle management to reduce recovery times and enhance data protection. These gains were enabled by adding a storage virtualization component to the company's backup solution and rewriting recovery procedures to match hot-site capabilities.

5.3 More reliable protection

A pharmaceutical company could not complete its daily backups within a 24-hour window and at one point had to do a restore from a 90-day-old backup tape. The company redesigned its storage architecture for higher capacity systems, added a tape hardware pooling solution and revamped its backup policies to allow for full weekly backups and incremental daily backups. This resulted in more reliable data protection, reduced backup windows and reduced manual handling of tapes.

5.4 Reduced costs, increased efficiency

Information lifecycle management approaches helped a large European insurance company move from labor-intensive reactive data protection processes to streamlined proactive processes. The effort resulted in reduced storage costs, increased storage utilization and enhanced data protection processes.

6 Recommended actions

- Thoroughly assess your data's protection needs and current data protection assets before making any infrastructure changes.
- Create a list of your most pressing data protection challenges. Prioritize these challenges in a manner that addresses the most critical issues first.
- Evaluate storage and backup resource management tools to globally monitor and manage your data protection infrastructure and processes.
- Gain business management support by briefing executive-level personnel on your organization's data protection challenges and potential opportunities.

7 Planning considerations

- Does your vendor have an established process for applying the principles of information lifecycle management to data protection?
- Does the assessment process address your entire data protection infrastructure and storage management processes or just a particular application?
- Does the information lifecycle management plan make optimum use of the most cost-effective storage systems that meet your data protection needs?

ABOUT STORAGETEK®

Storage Technology Corporation (NYSE: STK), a $2 billion worldwide company with headquarters in Louisville, CO, has been delivering a broad range of storage management solutions designed for IT professionals for over 30 years. StorageTek offers solutions that are easy to manage, integrate well with existing infrastructures and allow universal access to data across servers, media types and storage networks. StorageTek's practical and safe storage solutions for tape automation, disk storage systems and storage integration, coupled with a global services network, provide IT professionals with confidence and know-how to manage their entire storage management ecosystem today and in the future.

StorageTek products are available through a worldwide network. 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

MC 1043 A 04/04