WHITE PAPER

Optimizing storage capacity

Reduce storage costs by reserving your most expensive storage for your most critical data.

Data growth is skyrocketing and shows no signs of slowing down. To accommodate increasingly large data volumes, IT organizations typically add more primary disk capacity to their storage infrastructure. But this approach is costly and unsustainable over the long-term. Before investing in new disk capacity, IT managers should try to make their current infrastructure more efficient by optimizing existing capacity. This is accomplished by classifying data, eliminating unnecessary copies, moving less critical data to less expensive storage and repurposing allocated but unused capacity. Optimization can significantly reduce hardware and management costs.

JULY 2004

StorageTek®

1 Executive summary

- Adding more of the same disk capacity to accommodate data growth isn't a sustainable strategy.
- IT managers can delay or eliminate the need for new primary capacity by optimizing their existing infrastructure.
- The principles of enterprise information lifecycle management can provide a framework to assess data uses and storage assets, identify inefficiencies and adjust the infrastructure to maximize storage utilization.
- Drive out cost and complexity from the infrastructure by eliminating unnecessary and inefficient copies, moving less critical data to less expensive disk and recovering over-reserved capacity.
- Aligning storage costs with information's value can result in less dependence on expensive primary disk and a lower cost per online gigabyte.

2 Data growth requires IT to rethink storage strategies

In most industries, the rate of data growth is significantly outpacing increases in IT budgets and staff levels. Giga Research estimates that data volumes grew 40 percent to 75 percent during 2003.[1] Consider the following contributors to data growth:

- According to Enterprise Storage Group, fixed-content data like e-mail, presentations, reports and video files, is now growing at an annual rate of over 90 percent.[2]
- E-mail is a huge data management challenge at most companies. Storage Magazine estimates that a 3,000-person organization generates one terabyte of e-mail in one year.
- Many companies are proactively collecting and analyzing data about buying behaviors, which has caused customer transaction data to mushroom.
- Recent U.S. government regulations such as the Sarbanes-Oxley Act and the Health Insurance Portability and Accountability Act mandate multiple-year retention periods for some data.
- Virtually all businesses are affected by regulations regarding data archiving.

The mismatch between data growth and IT resources means conventional methods of treating the problem are not sustainable. Largely, those methods include adding more of the same expensive primary disk to the infrastructure, which is costly and increases the data management burden. A sustainable strategy built on the foundation of an enterprise information lifecycle management (EILM) approach is needed to effectively combat the problem. This approach provides a blueprint for optimizing valuable primary disk capacity while making better use of lower-cost storage classes.

3 Store data according to its value and purpose

IT managers can take practical steps today to manage data growth using the principles of EILM. EILM is a way to classify, manage and move information based on data's value and purpose at any point and time. In other words, it's about balancing the cost of managing information with its changing value. A large part of that cost is tied up in primary disk. An EILM strategy can reduce the need to add primary capacity and free unused capacity for application growth.

For example, a company's monthly financial reports may combine sales orders, shipments, inventory, expenses and other data. During the processing cycle, the finance department needs real-time access to verify and analyze the data. Therefore, it's logical to store the information on primary disk. After the reports are generated, the previous month's data is referenced less frequently as the focus changes to the current month. It then makes economic sense to offload the previous month's data from primary disk to lower-cost storage, like Serial Advanced Technology Architecture (SATA) disk or tape.

[1] "IT Trends 2003, Midyear Update: Storage Hardware," June 11, 2003, Giga Research.
[2] "Reference Information: The Next Wave," 2002, Enterprise Storage Group.

4 Optimize storage capacity

A simple three-step process can help IT organizations optimize storage capacity.

1. Assess the current storage environment.
2. Align data performance and availability requirements with storage class.
3. Refine and automate storage processes.

4.1 Step one: assess the current storage environment

An assessment of the current storage environment must include data valuation and an infrastructure audit. Data valuation starts with listing data types and ranking them based on how often the data is accessed, who uses it and other measures that make sense for a particular business. Once the data's value is classified, IT managers can map its current location to compare performance and availability requirements with storage class. This will likely reveal sub-optimal uses of capacity. Once these uses are identified and prioritized, IT managers can develop plans for recovering and reallocating primary disk capacity.

Causes of underutilized storage	Action
·· Non-critical data on high-performance disk	·· Assess use of data and map data to appropriate storage class
·· Data not referenced in past six months	·· Transfer to less expensive storage
·· Protection copies of less critical data	·· Transfer to less expensive storage
·· Storage allocated but unused	·· Review and re-justify allocations
·· Stranded capacity in decentralized disk systems	·· Consolidate into SANs where appropriate
·· Redundant copies (test copies, unnecessary duplication, no defined purpose)	·· Delete

Figure 1 is an example of a data map. Each dot represents a collection of data related to a specific application. The groups of data are stored in a range of storage classes, but are not aligned with the performance, availability or cost of each class. For example, disk mirrors, represented as two data points connected by dotted lines, are often stored on the same class of storage. Yet the protection copy probably has lower performance requirements.

In a lifecycle view of data, age is a key factor affecting its value. E-mail is a good example. Typically, access to e-mail messages falls off rapidly after a few days. There is virtually no access to messages more than a month old. If all e-mail is currently stored on primary disk, moving the older messages to secondary disk or tape will generate significant capacity and cost savings.

While age drives the usage frequency of some data, there are other variables that can influence performance requirements. The classification process helps to identify the variables and their relative weight. Objective measures like age and access are straightforward, but subjective measures should also be considered.

For example, the CFO's assistant e-mailed last year's earnings presentation to financial analysts. Six months have gone by and the presentation is outdated and rarely accessed. However, it should be readily available if requested by the CEO, shareholders, board members or anyone else. Therefore, the decision to continue storing it in the assistant's active e-mail folder is a wise one.

4.1.1 Assess disk utilization

After the data's value is classified, determine the current disk utilization in order to identify stranded capacity. For instance, direct-attached storage environments rarely exceed 50 percent utilization.[3] One reason for underutilization is that too much capacity is often reserved for future growth. Unused capacity is rarely identified and reclaimed.

Decentralized storage management also leads to inefficient utilization. A large number of non-networked or local network disks may collectively have more extra capacity reserved than necessary. Identifying and recovering this capacity is complicated because there isn't a single management view of disk resources. So more often than not, the capacity remains stranded.

Data protection and recovery strategies should also be targeted for disk utilization assessments. For example, the physical rotation of disk mirrors multiplies disk capacity. For every change on the primary disk volume, an identical change is made to a redundant backup volume. To get 24 hours of protection, a current copy could be rotated out every three hours and replaced by the oldest copy from 24 hours ago. This process requires duplicating the primary disk eight times for instantaneous recovery from hardware failure and limited corruption protection. Mirroring the data to lower cost disk may attain the same level of protection without using expensive primary disk space.

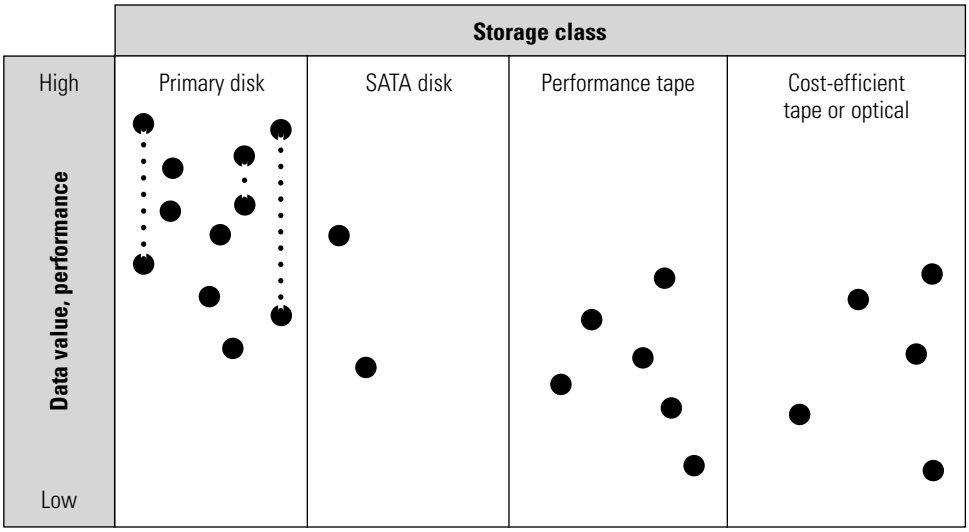

Figure 1. Mapping the value of data by storage system.

[3] "Improving Storage Utilization: The Benefits Are Often A Myth," January 23, 2003, Gartner Inc.

4.2 Step two: align data performance requirements with storage class

The assessment phase should yield a clear picture of data requirements and how well they are met in the current environment. Data of lesser value can be moved to lower cost, lower performance storage where it makes sense. There are a few ways to accomplish this: adjust the protection strategy; migrate data that is unused or unnecessary; and consolidate storage capacity.

4.2.1 Adjust the protection strategy

When optimizing storage capacity, use available technology to accomplish the same level of performance and protection at a lower cost. For example, mirrored physical copies may sometimes be replaced by mirrored logical copies. This technology stores only the data that has changed from one copy to the next, rather than creating a new complete physical copy. Primary disk capacity that had been used for physical mirror copies is now available for application growth.

Another example is the use of SATA disk as an alternative to primary disk. SATA disk arrays can store data at a fraction of the cost of high performance disk. Point-in-time copies often reside on expensive primary storage. These copies are used for both time shifting backup and for quick restore. SATA disk can often satisfy the performance needs for both of these activities.

4.2.2 Migrate data to other storage classes

Migrating unused content and unnecessary protection copies is another way to free storage capacity. Fixed-content data includes all types of non-changing electronic documents, like memos, presentations and medical records. This type of data is approaching 100 percent annual growth.[4] Research has shown that after three days, it's less than 50 percent probable that a piece of fixed-content data will ever be referenced again.[5] Yet much of this content is stored and protected on primary disk for long periods, eating up valuable capacity that could be put to more effective use.

Consider this example. A business has 16 terabytes of primary storage and disk mirroring in their environment. Eight terabytes is used for primary application data and eight terabytes for data protection. An analysis of the data showed a 50 percent drop in fixed-content usage after the first three days. The IT manager decided to move four terabytes of the application data that isn't being referenced from primary disk to SATA disk.

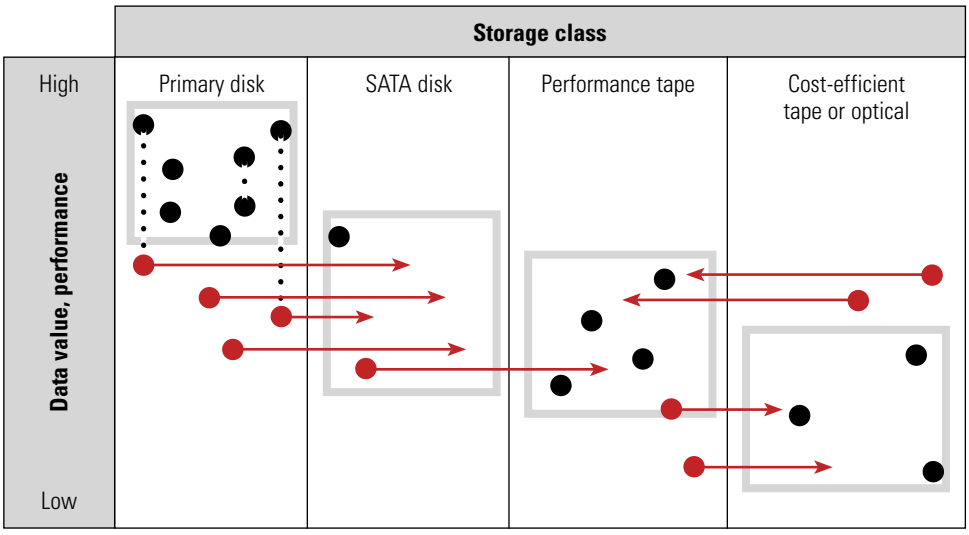

Figure 2. Migrating data to align performance requirements with storage.

[4] "Reference Information: The Next Wave," 2002, Enterprise Storage Group.
[5] Fred Moore, Horison Information Strategies, 2003.

Now factor data protection copies into the equation. A business impact analysis justified moving the protection copies off of primary disk. Using the same example as above, eight terabytes that were used for data protection are also moved to SATA disk. That leaves four terabytes of data on primary disk out of the original 16 terabytes. The 12 terabytes of recovered capacity are now available for other use. Assuming the purchase cost of SATA disk arrays is one-sixth the cost of high-performance disk, this process recovers substantial primary capacity value and lowers the overall cost per online gigabyte.

4.2.3 Consolidate storage and recover stranded capacity

Consolidating storage by class can simplify management and reduce disk capacity waste. In a decentralized environment, capacity is typically allocated by application. If the allocated capacity isn't used as planned, the extra capacity gets stranded within the network serving that application. By creating a pooled storage system, a smaller amount of total disk capacity can be reserved for application growth. This keeps reserved capacity to a minimum, and keeps it visible. It also simplifies management tasks by reducing complexity of the infrastructure. Highly efficient networked storage environments can operate at 70 percent to 90 percent capacity utilization, which is two to three times higher than typical utilization for direct-attached storage environments.[6]

4.3 Step three: refine and automate storage processes

Maintaining a balance between data performance needs and storage options is an ongoing process. As discussed earlier, IT departments can start the process today by classifying data, migrating data of lower value from primary storage to other storage classes and recovering stranded capacity. Automating these processes will further provide for the most efficient use of primary capacity. Policy-based data management tools can help by automatically migrating data as its value changes.

5 The opportunity

Balancing data value and storage performance can help free primary capacity and provide the following benefits:

- Accommodate growing data volumes
- Recover stranded capacity
- Improve storage management
- Lower the total cost per online gigabyte

Following are examples of companies that have successfully employed storage optimization techniques.

5.1 Transitioning data to lower priority storage

One large corporation found a way to balance availability and cost by automating payroll data management and migration. Since they consider payroll processing to be a mission-critical application, the data is stored on primary disk during the processing cycle.

After the pay cycle is complete, automated management tools move the data to midrange SATA disk arrays. Users can still access payroll data from the company's Web site for up to three months. After three months, the data is written to tape. For disaster recovery protection, the data is also replicated to a remote location.

5.2 Finding large amounts of unused data

In another example, a company found that 38 percent of the data it stored on primary disk had not been accessed for a year. By moving this data to lower-cost disk and tape, they opened up one terabyte of primary disk capacity without impacting performance. The IT department then implemented a rules-based process for archiving the data to secondary disk and then to tape.

[6] "Justifying Disk Storage Networking," January 20, 2003, Gartner Inc.

6 Recommended actions

- Consider using trusted experts to accelerate a capacity optimization plan.
- Start with a data assessment project. Identify what type of data is the environment, how it is used and where it is stored. Create a value hierarchy and map data to storage performance.
- Conduct an audit of primary disk utilization and identify unused capacity.
- Prioritize opportunities for freeing primary disk capacity and create an implementation plan that will yield the highest returns.
- Evaluate storage management tools to globally monitor and manage your primary storage infrastructure and processes.
- Measure success against cost per added gigabyte of primary data.

7 Planning considerations

- Does your storage vendor have an established process for applying the principles of enterprise information lifecycle management to data use?
- Does the assessment process address your entire primary storage infrastructure and management processes or just a particular application?
- Does the capacity optimization plan make best use of the most cost-effective storage systems while meeting data performance needs?

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

JT 0001 A 07/04