

The Storage Paradigm Shift: Surviving the Data Explosion with Disk- and Tape-Based Data Protection



The Storage Paradigm Shift: Surviving the Data Explosion with Disk- and Tape-Based Data Protection

By Stephanie Tilton

Executive Summary/Abstract

As data volumes explode, issues around bandwidth and resource constraints, storage costs, and backup windows make it difficult for organizations of all sizes to manage data backup and recovery efficiently and cost effectively. To ensure uninterrupted operations and growth, companies need a data-protection strategy that aligns with their business priorities, IT infrastructure, and regulatory requirements.

This white paper outlines the data-protection challenges experienced by a range of organizations and explains why tape-based storage alone may no longer be a sufficient option to resolve these issues. It then describes how the convergence of disk and tape—also known as disk-to-disk-to-tape (D2D2T)—along with deduplication technology enables a new solution that satisfies both short- and long-term data-protection needs.

The Crushing Impact of the Data Explosion

Data is the engine of organizations across every industry, powering the processes and solutions that enable innovation and a competitive edge. Fast and easy access to data is critical if businesses are to operate at their full potential. It's also necessary to avoid violating numerous regulations such as HIPAA and the Sarbanes-Oxley Act.

Yet organizations are grappling with a range of issues that make it complicated—and expensive—to ensure reliable access to data. One trend that threatens to crush existing data storage and backup strategies is the seemingly endless data explosion. According to IDC, the digital universe in 2011 will be 10 times the size that it was in 2006.¹ That's

¹ IDC, "The Diverse and Exploding Digital Universe: An Updated Forecast of Worldwide Information Growth Through 2011," 2008.

1,150 billion gigabytes of data, equivalent to 1.123 terabytes.

It's no wonder that mountains of data are piling up. Through the course of daily operations, businesses create multiple gigabytes of data that need to be backed up and accessed repeatedly. That's why data backup and recovery has become a top priority for businesses of every size.

But effective data backup and recovery is no small feat for the growing number of organizations operating in a distributed IT environment, with multiple data centers and remote offices/branch offices (ROBOs) around the world. According to Aberdeen Group, in such an environment companies struggle to:

- Ensure that someone is available and capable of managing data protection at remote sites
- Overcome backup issues and failures
- Find a sufficient backup window
- Manage costs associated with getting data offsite to the data center²

To complicate matters, as data growth soars, the cost of storing data also increases—especially when companies use legacy or under-performing backup technologies. Although the average cost per gigabyte of storage—whether disk-, tape-, or optical-based—is steadily declining, the savings are more than offset by the increasing rates of data growth seen in the past decade. Basically, storage capacity requirements are growing faster than the decline in storage costs. Using legacy or under-performing technologies for backup can worsen this problem

² Aberdeen Group.

because additional capacity must be added continually and more resources may be needed to guarantee a reliable backup process.

Organizations need a secure and efficient backup and recovery strategy to ensure optimal productivity and competitive advantage. Fortunately, revolutionary data technologies—and a paradigm shift in how backup technologies are used—promise to help organizations of all sizes overcome the challenges they face in managing and accessing their data.

Organizations and IT departments will be required to

- Transform the way that information is handled by the organization
- Implement organization-wide policies to deal with information governance (security, access, retention, and compliance)
- Carefully optimize storage to maximize security and retention but minimize cost

—IDC, *The Diverse and Exploding Digital Universe*

The Struggle to Protect Data

Companies across all industries depend on data to fuel their productivity and competitive edge, but data-protection issues vary by company size and setup. As can be expected, large enterprises must contend with huge data volumes. This makes it difficult to meet backup window requirements. Plus, ensuring fast restore times is no small task with so much data to manage.

To ensure uninterrupted operations, large enterprises have no choice but to provide reliable data backup and to scale without degrading performance. Complicating matters, they must effectively manage remote site data-protection requirements and satisfy a range of existing and potentially new regulatory requirements. And they

must accomplish all this while adhering to budget limitations.

Small-to-medium enterprises (SMEs) and ROBOs, on the other hand, face a different set of challenges, largely because they tend to prioritize PC and server upgrades over data protection due to budget constraints. In many cases, they use less expensive stand-alone storage devices rather than seemingly more expensive and complicated network-attached storage solutions. Although they may save a few dollars, the trade-off for many SMEs and ROBOs is the additional cost they incur from using a backup solution that requires more hands-on management. The dedicated IT staff necessary to start and complete backups are resources that could otherwise be focused on strategic IT activities and initiatives.

In addition, many of these organizations are unable to quickly recover data from offsite locations. This can lead to increased downtime, regulatory violations, an inability to meet SLAs, and loss of customer goodwill—all of which can hurt profitability.

60–70% of restores that occur two weeks after backup are for single files that have been lost or data sets that were corrupted.

—HP Research, Nearline Storage Division. 2008

A New View of Data Protection: Considering Short- and Long-Term Needs

To effectively address these challenges, all organizations need to reconsider their approach to data protection. Instead of thinking of it merely as backup, businesses need to understand how data protection helps address both short- and long-term organizational goals. Organizations need rapid, reliable data access to ensure business continuity and disaster recovery, but they must also retain data for long periods of time to satisfy compliance and archiving requirements.

In the *short term*, all organizations ideally want to make backups more efficient through an automated data-protection infrastructure. In addition, they need to enable quick recovery of lost or corrupted files or data sets.

Disaster Recovery Objectives

Recovery time objective (RTO) is the amount of time a business can tolerate a business process being down.

Recovery point objective (RPO) defines the acceptable amount of data a business can afford to lose.

To satisfy *long-term* objectives, organizations need to store data offsite. This enables them to protect against disasters and satisfy the data-retention requirements of specific regulatory agencies or industries. Plus, they need to be able to easily—and cost-effectively—scale to handle high rates of data growth.

What happens if organizations fail to address these challenges and needs? They will be less able to ensure nonstop business operations and efficient growth. In addition, they will likely struggle to support audits and satisfy regulatory requirements. As a result, they will incur significant costs—and damage to their reputations—as they experience downtime, productivity loss, and regulatory fines.

Tape Storage: The De Facto Standard for Data Protection

To understand why so many organizations find it difficult to overcome these challenges requires insight into traditional data-protection approaches. Tape technologies such as drives, autoloaders, and libraries have long been the de facto standard for data-protection storage. This is especially true for small and midsize businesses that choose tape-based data-protection solutions over other, seemingly more expensive, technologies.

Tape storage has several distinct advantages over other storage mediums, including the following:

- **Easy portability:** Data can be taken offsite for disaster-recovery protection.
- **Long shelf life:** Data can be archived for up to 30 years, meeting regulatory data-retention requirements around the world.
- **Low cost:** The price of tape is usually about half that of disk-based solutions when the costs of hardware (including tape media) are spread across multiple terabytes of data.
- **Density:** Large amounts of data can be stored on a very small and inexpensive data cartridge because, with compression, cartridge capacity can exceed a terabyte.

However, tape storage alone may not provide enough protection for business data. After all, most organizations only consider the immediate IT environment to determine how much backup capacity is required and how quickly backups need to be completed. Equally important is the ability to recover data rapidly and reliably. With this in mind, some organizations turn to disk to ensure fast backups and recovery. The problem is that these organizations are compromising by making a choice between disk and tape.

The Solution: Disk-to-Disk-to-Tape with Deduplication

Fortunately, organizations can take advantage of the convergence of disk and tape—also known as disk-to-disk-to-tape (D2D2T)—to better handle data protection. D2D2T technology leverages the best features of both disk and tape storage to create a cost-effective and comprehensive data-protection solution that can meet both short- and long-term needs.

Disk-based storage works with the backup application to schedule and automate daily backup jobs, requiring fewer IT resources (and less human intervention) to manage this task. Because most D2D2T solutions are network-attached, backup jobs

from multiple servers can be consolidated onto a single disk (or caching device) for easier data and device management. Backup to disk also enables fast data searches and quick data recovery, minimizing downtime and required recovery efforts. After data has been backed up to disk, it can then be selectively copied to tape storage for offsite disaster-recovery protection and to meet various regulatory agencies' and industries' long-term data-retention requirements.

Additionally, some D2D2T solutions support replication of backup jobs between remote sites and data centers. This can help to further minimize the need for local IT intervention in the data-protection process. It also enables additional consolidation in the data center, where multiple remote sites can share longer-term tape storage resources.

A critical technology behind the success of the D2D2T approach is data *deduplication*, a process by which redundant data—in other words, data that has been previously backed up—is eliminated from future backup jobs. This allows businesses to keep more data online for longer periods of time—at significantly lower costs—enhancing the ability to provide fast recovery from disk.

Although multiple types of data deduplication technologies are now available, data deduplication that occurs at the *block* level is considered the most efficient, providing deduplication ratios as high as 50:1.

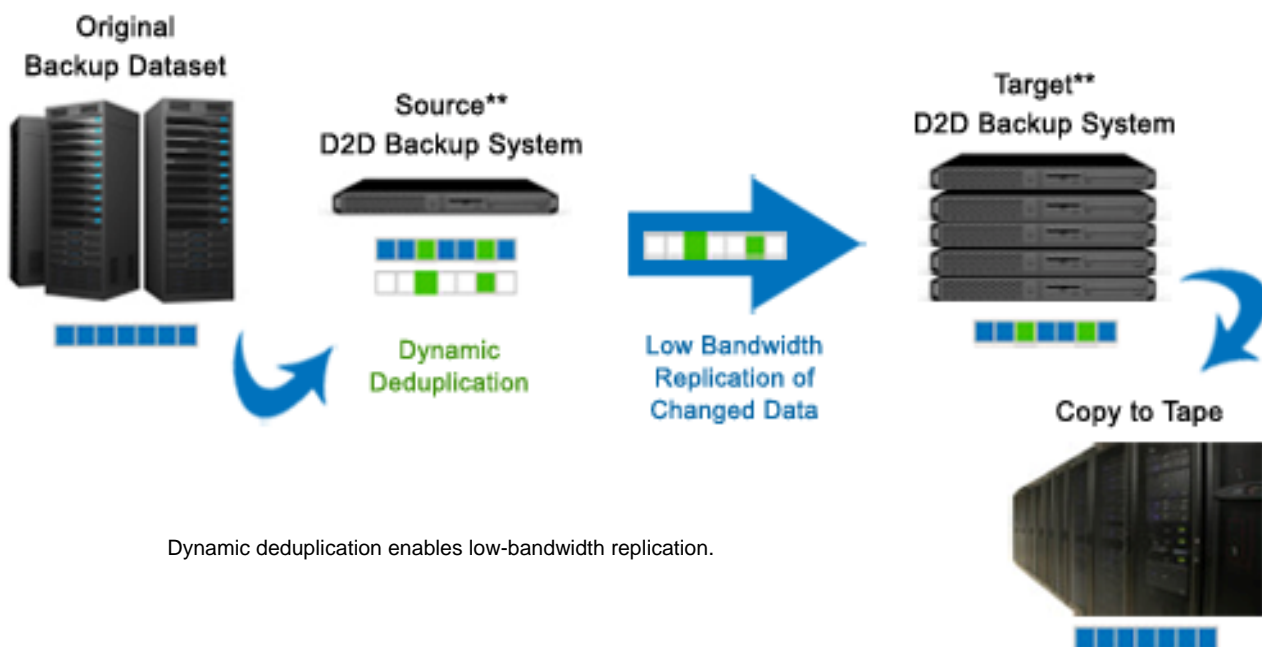
How Does Deduplication Work?

Deduplication technology looks at data on a sub-file or block level and determines if it has seen the data

Disk-based backup systems can enhance data protection for a business of any size. Integrating deduplication technologies allows customers to leverage up to 50 times more capacity to ensure their business has faster and reliable access to stored data.

—Dave Russell, Vice-President, Gartner Inc.

before. If it hasn't, it stores the data. If it has seen the data before, it ensures that the data is stored only once and merely points to all other references of the data. This process makes it viable to replicate backup data over low-bandwidth WAN links—for off-site protection—because only changed data is sent across the connection to a second device. For remote sites that have little or no IT support, low-bandwidth replication can provide a simple, low-cost,



Dynamic deduplication enables low-bandwidth replication.

efficient method of storing data offsite to protect against site-wide disasters.

The benefits of deduplication are numerous:

- Makes remote backup and disaster recovery more affordable for a wider range of businesses
- Optimizes the use of disk storage capacity during backups, and minimizes add-on storage requirements
- Minimizes the bandwidth needed to transfer backup data to offsite archives
- Lowers dependence on remote tape-based backups
- Enables quick data recovery in case of an outage
- Reduces the overhead required for backup, media management, and off-site vaulting

What to Look for in a D2D2T Solution

A number of D2D2T solutions are now on the market, but feature sets, deduplication/replication technology, and price points vary tremendously. With that in mind, organizations should look for a solution that satisfies the following requirements at a minimum:

- **Integrates with the existing hardware, software, and network environment.** Some solutions require organizations to purchase additional software and/or hardware, defeating the purpose of minimizing Total Cost of Ownership (TCO). The ideal solution integrates with the existing environment, requiring no additional capital expenditures or management overhead.
- **Is optimized for backup.** Although many vendors offer appliance-based solutions, most of them have been retrofitted to support backup. As a result, they can be challenging to deploy and maintain. The ideal solution was designed for backup, incorporating features that simplify setup, maintenance, and management.

- **Supports multiple deduplication methods.** Most vendors offer only one deduplication method, but no single method is best in all circumstances. In enterprise environments, the ideal solution supports scalability with performance, enabling users to decide what data types are deduplicated and to reduce the deduplication load if it is not yielding the expected or desired results. In SME and ROBO environments, the best solution is cost-efficient.
- **Satisfies a range of data-protection needs.** Many solutions are designed to address a narrow backup requirement. However, data protection is about more than backup. The ideal solution enables organizations to address both short- and long-term data-protection needs.

Why HP Is the Best Choice for Data Protection

Combining the best features of both disk-based and tape-based storage, HP offers the broadest range of D2D2T data-protection solutions in the industry, providing affordable, reliable, and easy-to-use solutions for businesses of all sizes. With its portfolio of StorageWorks D2D Backup Systems, Virtual Library Systems (VLS), and LTO Ultrium tape solutions, HP D2D2T affords a number of benefits through its data-protection solutions, including the following:

- Ensures low TCO
- Designed specifically for backup
- Accommodates various data-protection environments
- Supports low-bandwidth data replication
- Offers seamless support for both disk and tape

Ensures Low TCO

This affordable family of data-protection solutions includes a number of features that help ensure low TCO. Seamless integration into existing IT environments eliminates hidden upgrade costs and reduces the need for additional components that are typically required to enable a D2D2T solution. These solutions support direct connect to either Ethernet or

Fibre Channel networks—utilizing current backup applications and software licenses—and work with most server platforms and operating systems. Management from a central data center using an Internet-based GUI reduces the need for remote IT resources. The wizard-based installation requires just three simple steps and is typically complete in less than an hour. Plus, data-deduplication technology enhances storage efficiency and allows low-bandwidth replication—both of which contribute to cost savings.

Designed Specifically for Backup

HP has created its backup appliances from the ground up to address specific data-protection needs. The disk-based nature of the appliance enables automated backup for multiple servers and fast access to data for recovery. Automated replication further reduces the need for human intervention, and an intuitive web-based browser interface enables anywhere, anytime management so IT staff can view results or change settings. And because the solution is easy to install, manage, and maintain, no IT training is required.

Accommodates Various Data-Protection Environments

Recognizing that each organization's requirements differ, HP offers a choice of deduplication technologies. Small and mid-size companies, as well as remote and branch offices, can take advantage of HP Dynamic deduplication that offers the best technology footprint for deduplication at an affordable price point. HP supports Dynamic deduplication on its entire line of HP D2D Backup Systems as a standard feature that requires no additional licenses or upgrades. The deduplication process occurs at the time of the backup (in other words, as inline deduplication) and is compatible with all mainstream backup applications.

Enterprise and Fortune 500 data centers benefit from HP Accelerated deduplication, optimized for scalability and backup performance, which is offered through the HP StorageWorks family of VLS. The Accelerated deduplication option leverages post-

processing technology, which deduplicates data after backup jobs are completed. This results in higher-performance backups, immediate access to data for recovery at all times, and a scalable architecture that doesn't slow down as the IT environment grows.

Supports Low-Bandwidth Data Replication

Support for low-bandwidth data replication enables simple and affordable replication of backup data to a central offsite facility or data center. Because data can be replicated back to a central site, organizations can reduce or eliminate the use of removal media, such as tape, at remote sites. Plus, backup data can be consolidated at the central site, where it can be copied to tape for long-term archiving.

Offers Seamless Support for Disk and Tape

Tape drives are still a critical component of a robust data-protection strategy, providing the most cost-effective way for copies of data to be stored offsite for disaster recovery or long-term archiving. The HP D2D Backup System works with existing tape drives and backup software to migrate data to tape for offsite storage. Alternatively, the HP D2D Backup System models are available with integrated LTO-2 or LTO-3 Half-height SAS tape drives. A tape-offload feature allows organizations to copy or export cartridges from the D2D Backup System to a physical tape without slowing down the network.

Take Advantage of the Storage Paradigm Shift

With so much value locked up inside information, organizations need to employ an effective backup and recovery strategy. Unfortunately, businesses of all sizes struggle to access and protect their data for a variety of reasons, leaving them at risk for business interruptions, negative publicity, and regulatory violations.

As data grows at an alarming rate, organizations cannot afford to think about data protection in isolation. It's no longer acceptable to compromise by choosing between tape-based protection for cost-

effective offsite storage and disk for fast backups and recovery. To effectively address both short- and long-term needs, businesses must adopt a new way of thinking about data protection. In short, organizations need a secure and efficient backup and recovery strategy and infrastructure to ensure optimal productivity and competitive advantage.

Fortunately, a paradigm shift in how tape- and disk-based backup technologies are used promises a new way. HP offers affordable D2D2T solutions that leverage the best features of both disk and tape storage—along with deduplication technology—to deliver the access and recovery speed of disk with the long-term protection benefits of tape. As a result, organizations can cost-effectively and efficiently store data while ensuring optimal performance, better business efficiency, and compliance with global regulatory requirements.

To find out how HP can help you address your short- and long-term data-protection needs, visit:

www.hp.com/go/d2d

www.hp.com/go/d2d2t

www.hp.com/go/vls

www.hp.com/go/deduplication

HP

HP is a technology company that operates in more than 170 countries around the world. We explore how technology and services can help people and companies address their problems and challenges, and realize their possibilities, aspirations and dreams. We apply new thinking and ideas to create more simple, valuable and trusted experiences with technology, continuously improving the way our customers live and work. No other company offers as complete a technology product portfolio as HP. We provide infrastructure and business offerings that span from handheld devices to some of the world's most powerful supercomputer installations. We offer consumers a wide range of products and services from digital photography to digital entertainment and from computing to home printing. This comprehensive portfolio helps us match the right products, services and solutions to our customers' specific needs. To learn more, visit <http://www.hp.com>.

IT Business Edge

IT Business Edge delivers the information, analysis and context business technology decision makers need to maximize returns on IT investments and align IT initiatives with business objectives. As a Technology Intelligence Agent, IT Business Edge provides content different from that of a traditional IT publisher, news service or analyst firm. Our editors monitor all these sources—plus many others—for critical IT information that they translate into actionable advice for high-level IT and business managers. Subscribers access our practical content and useful decision-making tools through a rich Web site, targeted e-mail newsletters and varied RSS feeds. All these outlets feature our business-focused blogs, exclusive interviews with field experts and industry insiders, plus our database of more than 20,000 abstracts summarizing content from 2,500-plus sources. Visit our Web site at: <http://www.itbusinessedge.com>.

About the Author

Stephanie Tilton is a B2B marketing consultant who helps clients advance the buying cycle with engaging content. She has written over 100 white papers and case studies for start-ups and leading technology brands, and is the author of the eBook *5 Steps to a White Paper that Pulls in the Perfect Prospect*. Her best-practice articles have appeared in *productmarketing.com* magazine, and on sites including the American Marketing Association, BNET, and SalesVantage. To find out how Stephanie can help you educate prospects, promote your thought leadership, and ratchet up the results of lead-generation campaigns, visit <http://www.tentonmarketing.com> or contact her at stilton@tentonmarketing.com.