

# White Paper

---

## **Storage-efficient Data Protection and Retention**

*By Lauren Whitehouse*

**April, 2011**

---

This ESG White Paper was commissioned by IBM  
and is distributed under license from ESG.

## Contents

Introduction .....	3
Data Growth’s Impact on Data Protection and Retention .....	3
Cost .....	3
Risk .....	3
The Dilemma: Minimizing Risk versus Minimizing Cost .....	4
IBM Portfolio Solutions Addressing Data Protection and Retention .....	6
Snapshot Technology .....	6
Disk and Tape Backup .....	6
Replication .....	6
Archiving .....	7
The Bigger Truth .....	8

This document was developed with IBM funding. Although the document may utilize publicly available material from various vendors, including IBM, it does not necessarily reflect the positions of such vendors on the issues addressed in this document.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188.

## Introduction

The volume of data managed by organizations today is growing exponentially as a result of increased dependence on digital content and long-term data retention. While businesses can become more agile via the use of information technology, there's a price to pay: conducting business requires uninterrupted access to data and systems. With limited tolerance for downtime, organizations adopt contingency plans that include auxiliary systems and copies of data. This limits risk, but it also increases costs. Since risk mitigation and business agility cannot be compromised, IT organizations seek data protection solutions that optimize their storage investments. [IBM](#) delivers a portfolio of data protection and retention solutions that can improve efficiency in many ways. This paper will review the challenges introduced by data growth and focus on IBM data protection and retention products/technologies that address storage efficiency.

## Data Growth's Impact on Data Protection and Retention

Data growth shows no signs of abating: ESG's total worldwide forecast through 2015 for archive capacity predicts a 56% compound annual growth rate (CAGR).<sup>1</sup>

Companies of all sizes are ever-dependent on digital content to conduct business. The increased value of information causes companies to store more data and keep it longer. It's not only e-mail data or databases supporting content management, financial, sales order processing, supply chain management, and customer relationship management applications; it's the various files, or "unstructured" data, growing at a blistering pace. In fact, file data made up the lion's share (76%) of archived data in 2010.<sup>2</sup>

Data growth is driven by several factors. Of course, newly created data is a contributor to information's expanding footprint in business. However, the increasing volume of data is also a result of rich media types, extended retention driven by corporate or regulatory compliance, or just plain apathy regarding data disposal. There are also numerous copies of data made for data protection and other purposes that can compound data growth.

### Cost

The costs of managing multiple copies of data can be high. With each phase of data growth, ongoing investments in storage capacity, compute power, and network bandwidth are required. The impact on operational expense (OpEx) is seen in data center floor space requirements and the costs incurred to support incremental power and cooling, operational staff, bandwidth, portable media, and eDiscovery.

There are other consequences of this information explosion. Growing storage environments can become more complex, lead to unpredictable resource utilization, and increase off-site record retention costs. Failure to meet service level agreements and regulatory mandates for data protection or retention can lead to non-compliance fines.

### Risk

Risk may also be introduced if data protection strategies and systems cannot keep pace. ESG research found that over 70% of organizations can experience only three hours or less of downtime for tier-1 data before suffering adverse business impact (see Figure 1).<sup>3</sup> System or application downtime introduces higher risk for missed business opportunities, employee or customer dissatisfaction, lost data, lost revenue, damage to the company's reputation, lost productivity, and even legal liability.

Businesses have to be responsive and resilient for any eventuality—seamlessly taking advantage of opportunities while minimizing threats. Business resiliency is buoyed via data protection solutions.

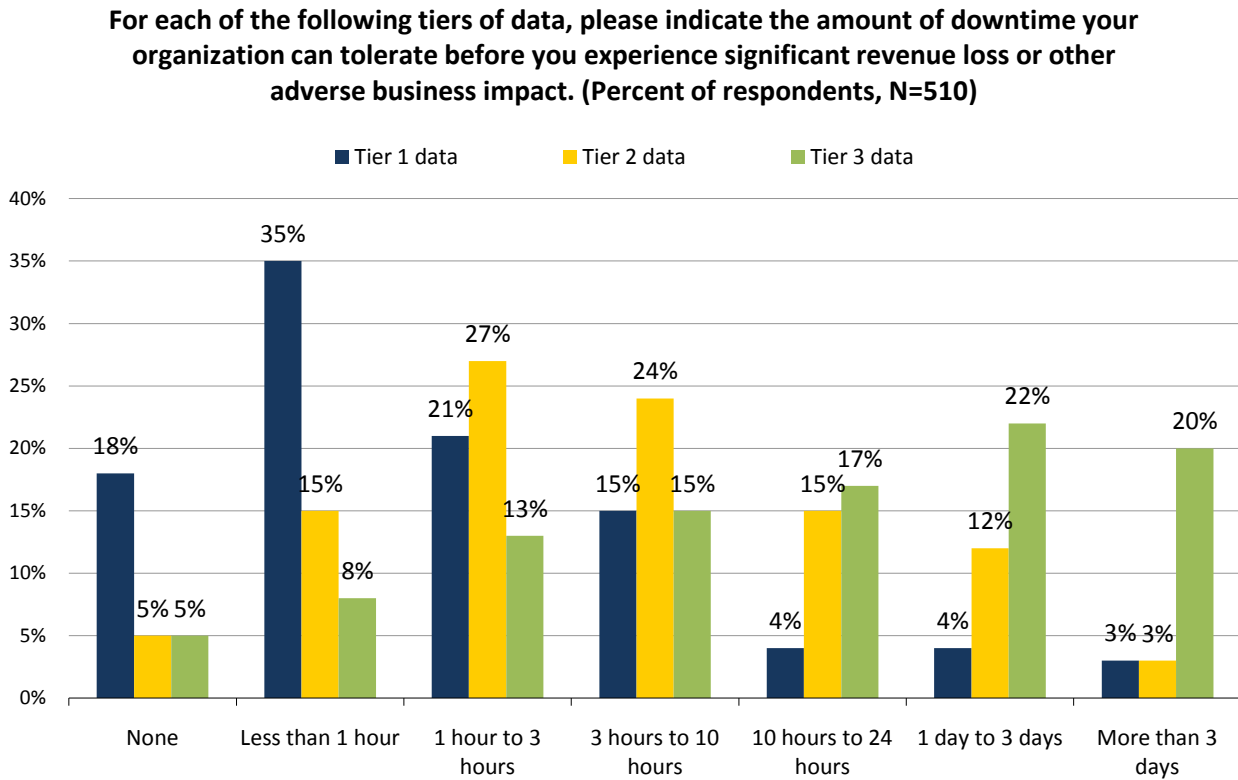
---

<sup>1</sup> Source: ESG Research Report, [Digital Archive Market Forecast 2010-2015](#), June 2010.

<sup>2</sup> Ibid.

<sup>3</sup> Source: ESG Research Report, [2010 Data Protection Trends](#), April 2010.

Figure 1. Downtime Tolerance By Application Tier



Source: Enterprise Strategy Group, 2011.

## The Dilemma: Minimizing Risk versus Minimizing Cost

Companies are often faced with the challenge of investment choices: minimize risk or reduce costs. This dilemma is especially perverse since backup copies created for operational and disaster recovery compound the data growth problem—and the tumbleweed effect more data has on the environment and costs. What if you could save money on data protection without increasing risks and redirect funds to IT investments that help grow the business?

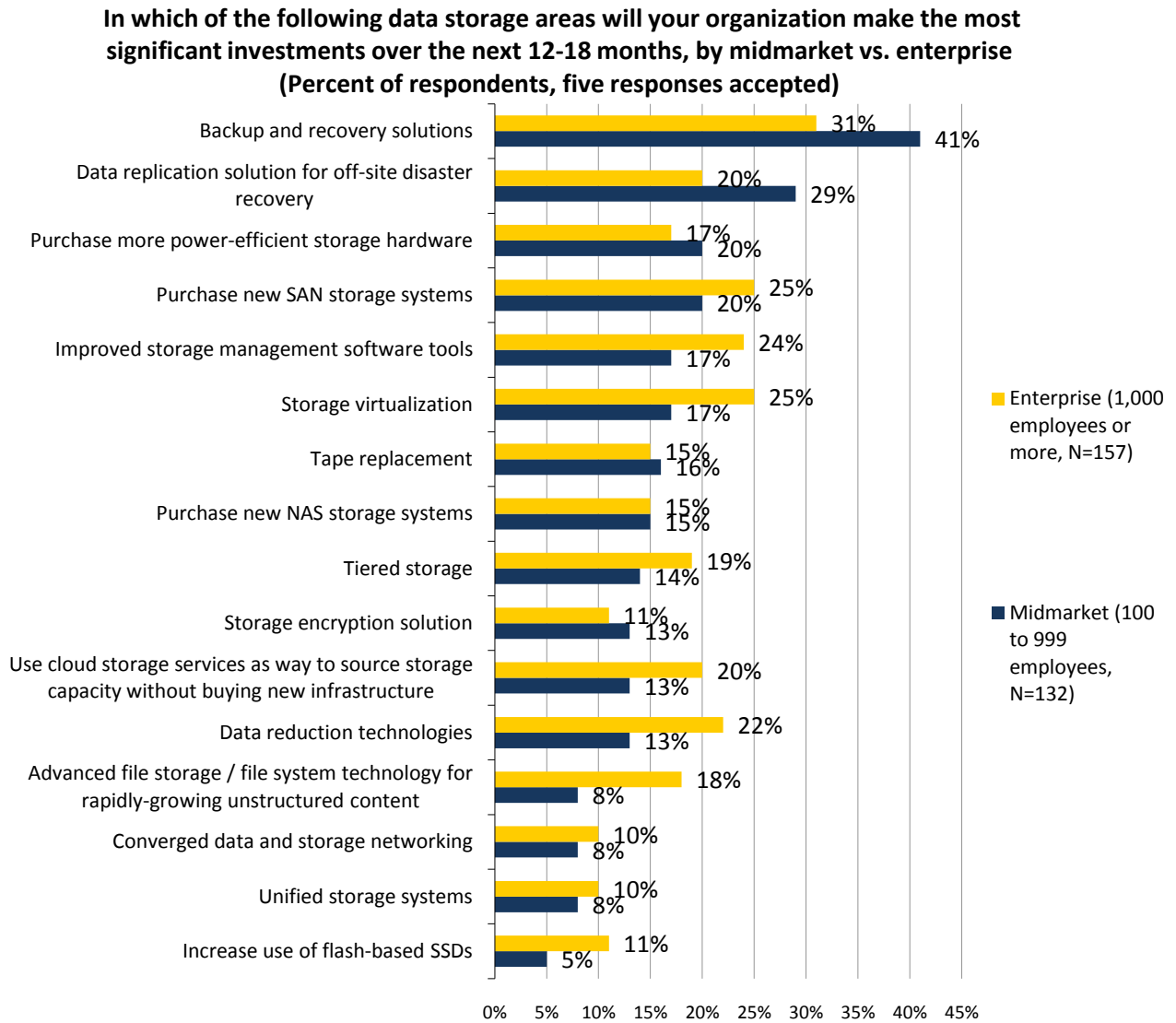
That’s what IBM does today with its data protection and retention solutions; they increase efficiency and help reduce costs without making concessions on risk.

Backup copies are created to recover systems and data. Data protection technologies such as backup, snapshot, and replication make a secondary copy of a given data set; in the event of the loss of the primary copy of that data, business operations may continue. IBM offers its flagship Tivoli Storage Manager (TSM) backup software, TSM FastBack continuous data protection (CDP), mirroring solutions, and cloud-based backup. IBM solutions integrate with tape and disk storage, providing an easy way to control costs over time. Further, IBM SmartCloud data backup services deliver scalable, managed data protection services on- or off-premises for servers and PCs.

Introducing efficiency in backup and recovery processes is one approach to cost savings. ESG research found that IT organizations will invest in driving storage efficiency—right behind mitigating risk with backup and recovery. As shown in Figure 2, IT plans to drive storage efficiency with investments in storage virtualization, power efficient storage, data reduction, and tiered storage and is prioritizing the use of replication for disaster recovery.<sup>4</sup>

<sup>4</sup> Source: ESG Research Report, [2011 IT Spending Intentions](#), January 2011.

Figure 2. Top Storage-Related Investment Initiatives



Source: Enterprise Strategy Group, 2011.

The efficient capture of data is one aspect of streamlining processes. For example, enabling rapid backup copies via snapshot technology accelerates backup and recovery operations. Leveraging disk storage systems as backup targets speeds backup and, importantly, recovery. Providing incremental or delta block capture or applying deduplication technology to backup copies can significantly decrease the volume of data copied and stored. This level of efficiency can be seen in TSM and TSM FastBack software and in ProtecTIER deduplication disk storage as well as in the snapshot and replication solutions IBM offers.

Similarly, efficiency techniques are available in the transfer of data via IP replication, mirroring, and tape media management features of IBM storage systems. As an alternative to disk, tape media delivers portability and value for both long-term retention and off-site storage for disaster recovery. Tape systems can store aged backups for a fraction of the cost of disk, so they can improve the total cost of ownership (TCO) of backup environments.

Another is implementing a dedicated archive solution to reduce the volume of data in the primary storage environment. This not only improves the performance and effectiveness of production systems, it also results in lower secondary (backup) storage capacity requirement, as well as reduced time to complete backup and recovery jobs. Lastly, an effective archiving solution can limit exposure to legal and regulatory actions by helping to meet corporate and regulatory mandates for retention and discovery.

## IBM Portfolio Solutions Addressing Data Protection and Retention

As a result of data growth, limited tolerance for downtime, and budget constraints; data protection and retention processes have to be faster, smarter, and cheaper. IBM's data protection and retention portfolio has expanded over the years to meet customer requirements for more speed and efficiency.

### Snapshot Technology

Snapshots are replicas of production data. The copy operation has minimal impact on production system performance and the restoration operation offers near-instant recovery. Leveraging this technology for application-specific, transaction-consistent protection provides application-aware snapshots that aid in faster application recovery. Snapshots that only capture and transfer changed blocks accelerate processes and minimize capacity requirements. They also reduce the risk of data loss since more frequent captures can be performed.

Integrating snapshots into the data protection environment is simplified with IBM TSM. Policy settings in TSM can trigger snapshots and snapshot management can streamline recovery processes. Migrating snapshots off site can be facilitated using TSM integration.

Tivoli Storage FlashCopy Manager integrates with IBM System Storage DS8000, SAN Volume Controller (SVC), Storwize V7000, and IBM XIV storage systems. It also protects Microsoft applications on DS3000, DS4000, and DS5000 storage systems via Microsoft Volume ShadowCopy Services (VSS).

### Disk and Tape Backup

IBM TSM protects, retains, and ensures accessibility to data. Its architecture and design distinguishes it from comparable solutions with features such as disk-to-disk backup, progressive incremental backup policies, and unified recovery management. These and other features are proof points of the level of optimization in the TSM platform.

IBM ProtecTIER offers an alternative to disk and tape backups, enabling backup and recovery performance improvements and significant cost reductions. ProtecTIER features inline data deduplication, which optimizes storage capacity, and its native IP replication technology facilitates off-site storage of backup copies to facilitate disaster recovery.

IBM Business Continuity and Resiliency Services offer public, private, and hybrid cloud solutions for backup and recovery of data, wherever it's located. In addition to deduplication and scalability, it offers options for continuous or scheduled backups, and managed services for inside or outside the data center. With over 160 resiliency centers worldwide, IBM can deliver a single-vendor solution to protect data enterprise-wide.

### Replication

For aggressive recovery time objectives, data replication becomes an essential technology. Synchronous or asynchronous replication of data to one or more secondary storage systems can provide an exact mirror copy of primary data to another storage system. In the event of a system failure, the mirror copy can be used for rapid fail-over.

Replication technology captures a data changes with minimal overhead. The IBM data protection and retention portfolio offers multiple replication options from storage system features to software. For example, Metro Mirror delivers a synchronous copy option that continually updates a secondary copy of a volume to match changes made to a source volume in the same or separate storage devices. IBM Global Mirror accomplishes the same, but asynchronously and over longer distances. IBM XIV offers built-in synchronous or asynchronous mirroring with customizable recovery points. IBM SAN Volume Controller can mirror virtual disks up to 10 km, and can integrate with VMware vMotion to enable transparent application migration between data centers.

IP Replication takes advantage of in-place IP networks, reducing costs by using existing network infrastructure. Providing fast, real-time data synchronization of storage volumes eliminates the distance limitations of other solutions and provides the ability to recover systems rapidly. IP replication is available in both the mainframe virtual tape library (VTL) solution, IBM TS7700, and the open systems VTL, ProtecTIER, as well as software such as TSM FastBack.

## **Archiving**

Archiving solutions facilitate the movement of data from a production environment to a separate one. The efficiency of archiving is derived from features like deduplication (for storage optimization), indexing (for searchability), and retention management. IBM's archive strategy involves a series of hardware, software, and services offerings.

IBM Information Archive is a unified storage repository for archived information. The disk storage appliance also supports IBM and non-IBM tape systems, enabling a scalable combination of disk and tape for archive storage needs. By supporting disk and tape and automating the movement of data between the two, it controls storage costs which can impact both CapEx and OpEx savings.

Information is often archived for compliance and electronic discovery purposes: information relevant to a specific legal or regulatory matter must be preserved until the matter is resolved, or corporate policies prevent the deletion of certain information in case it is needed to support or defend a claim. IBM eDiscovery Analyzer provides investigators, auditors, and legal professionals with a quick and easy way to find relationships in data using advanced analytics and visualization.

IBM tape storage solutions offer the latest advancements in performance and usability, providing an affordable option for archive workloads. For example, IBM Linear Tape File System (LTFS) works in conjunction with LTO5 tapes and libraries; making data on tape as easy to access as data on disk.

IBM Enterprise Archive Services combines IBM tools and practitioner experiences to improve compliance and reduce risk and costs. The service involves the assessment, design, and implementation of an information archive strategy.

## The Bigger Truth

IT organizations are under increased pressure to reduce risks and costs, particularly in the areas of data protection and retention. At the same time, there is an increased dependence on data to run the business, including the need to manage constant data growth and to make sure data is available when requested. IBM has a portfolio of solutions to ensure that data is safe and accessible while also introducing efficiency to deliver savings.

IBM solutions and services can help organizations with the “minimize-risk/minimize-cost” challenge. Its data protection portfolio provides a multi-layer defense strategy and streamlines recovery to mitigate risk while also driving down the time it takes to conduct backup and recovery and reducing the cost of storing backup copies.

IBM’s retention/archiving solutions amplify cost savings and risk mitigation by enabling long-term data to be maintained in an easy-to-search repository. Benefits of an effective archiving strategy include improved access to information, reduced legal exposure, and improved efficiency and performance of both production and backup systems.





Enterprise Strategy Group | **Getting to the bigger truth.**