THE COMING STORAGE ARMAGEDDON:
WHAT YOUR DATABASE ADMINISTRATOR ISN'T TELLING YOU

An Executive Brief Sponsored by IBM

March 2011

THE COMING STORAGE ARMAGEDDON: WHAT YOUR DATABASE ADMINISTRATOR ISN'T TELLING YOU

EXECUTIVE SUMMARY

If your company is like most, you probably believe you have a good handle on storage growth. Sure, the storage budget increases every year, but you can easily rationalize that. Storage growth means that your company adheres to best practices for data retention and backup; that you've automated your business processes; that you're keeping up with new compliance regulations. Besides, storage is cheap now compared to when you started in the industry. So there's no problem, right?

Wrong.

The truth is that data growth is rapidly outpacing our ability to process, analyze and store the data we're collecting. In fact, *enterprise storage needs are increasing by up to 40 percent per year.* That means up to 40 percent more equipment; 40 percent more rack space; 40 percent more people; 40 percent more power; 40 percent more budget dollars. Every year.

And the problem goes well beyond resources consumed. Unbridled data growth combined with inefficient storage will lead to application performance problems—which can cause productivity loss, customer dissatisfaction, and hits to the bottom line.

Smart IT leaders are not waiting until the torrent of data shuts down key operations. Instead, they are scrutinizing their current processes, and weeding out those that are clearly inefficient or not scalable. They're redesigning their storage approach based on three simple tenets:

- **Reduce the amount of data to be stored.** Common technologies like deduplication and compression can reduce the volume of stored data by 50 to 80 percent, and can improve application performance in the process.

- **Store data in the right place.** Data that's accessed most often should be right at hand; less commonly used data should be tucked away in a less costly storage solution.

- **Optimize the existing storage environment.** Before you buy more, make sure you're fully utilizing the equipment you already have.

In this white paper, Stratecast examines the business risks associated with a "let it be" approach to data storage. We present surprisingly easy and cost-effective solutions that can have a high impact on improving the storage environment. Finally, we look at how IBM's suite of storage solutions can introduce scalability and efficiency to the storage environment—staving off storage Armageddon.

INTRODUCTION

Do you remember the classic *I Love Lucy* episode about the candy factory? At first, Lucy and Ethel easily handle the chocolates moving down the conveyor belt, transferring them safely to appropriate boxes. But, as the belt speeds up and the volume increases, they are unable to keep pace. Chaos ensues, productivity is halted, and (to the horror of chocolate lovers everywhere) the precious bonbons are ruined.

Can't happen in your data center? Think again. We grant that your database administrators are more skilled (and coordinated) than TV's madcap duo. But a storage environment established to handle a relative trickle of data is not equipped for a sudden, dramatic, and continuing onslaught. *The "new normal" is a never-ebbing torrent of data. If you don't prepare now to handle it, chaos will ensue, productivity will halt, and your valuable data will be unusable.*

How Bad Is It, Really?

Armageddon? Torrent? Perhaps you suspect that we are being alarmist in the way we describe what you perceive as a gentle increase in data.

But let's take a look at some jarring industry statistics and facts:
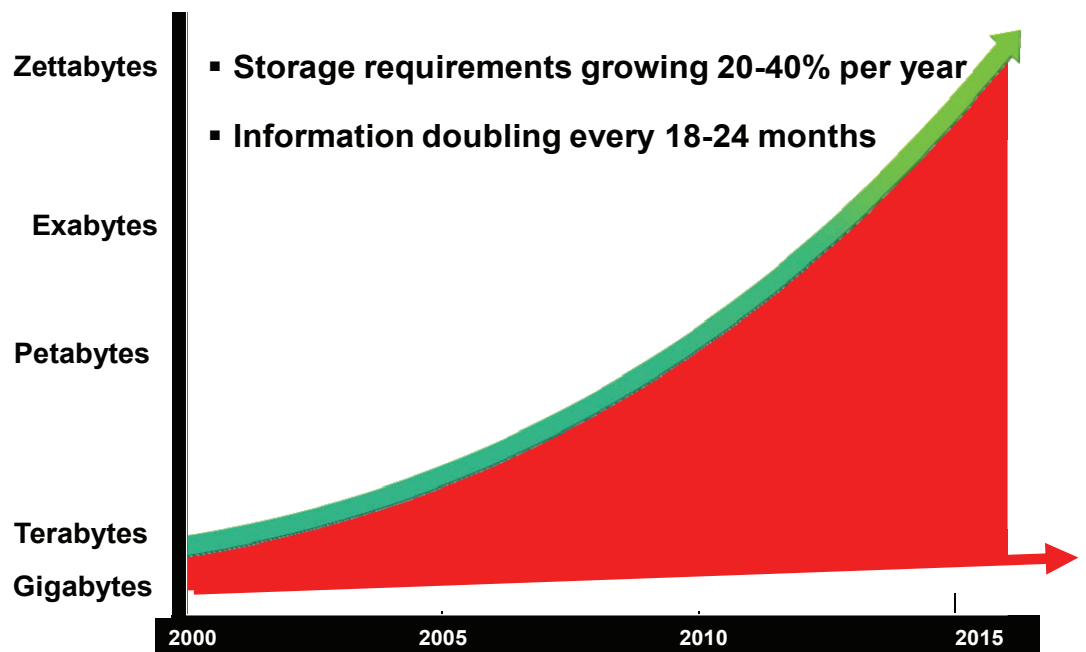
- *Worldwide data is doubling every 18-36 months, according to most industry prognosticators.* Total current data is measured in thousands of Exabytes (EB).

- By 2013, *the amount of data flowing over the Internet will exceed 667 Petabytes (PB) annually.*[1]

- *87 percent of enterprises store more than 1 Terabyte (TB) of data*, up from 74 percent in 2009. *29 percent of enterprises store more than 100 TB*, up from 24 percent in 2009.[2]

1 Zettabyte (ZB)	= 1,000,000,000,000,000,000,000 Bytes = 1 trillion Gigabytes = 1 billion Terabytes = 1 million Petabytes = 1024 Exabytes
1 Exabyte (EB)	= 1,000,000,000,000,000,000 Bytes = 1 billion Gigabytes = 1 million Terabytes = 1024 Petabytes
1 Petabyte (PB)	= 1,000,000,000,000,000 Bytes = 1 million Gigabytes = 1024 Terabytes
1 Terabyte (TB)	= 1,000,000,000,000 Bytes = 1024 Gigabytes

[1] "Data Data Everywhere," The Economist, February 25, 2011, http://www.economist.com/node/15557443

[2] 2010 State of Enterprise Storage Survey, Infoweek, Jan. 31, 2010, http://www.informationweek.com/news/storage/systems/showArticle.jhtml?articleID=222600268

March 2011

Figure 1 - Estimated Data Growth - Worldwide

Zettabytes

- **Storage requirements growing 20-40% per year**
- **Information doubling every 18-24 months**

Exabytes

Petabytes

Terabytes
Gigabytes

2000   2005   2010   2015

Source: IBM

The hockey stick growth reflected in Figure 1 isn't restricted to data-intensive global firms, like Facebook with its 40 billion stored photographs. *In fact, the average annual rate of data growth across all enterprises is 20 to 40 percent*, driven by a few key business trends:

- *We are automating more business processes*, and collecting and analyzing more data. Many applications that didn't exist a few years ago (e.g., telemetry, telehealth, utility monitoring) are now common. Furthermore, companies are storing their data longer to provide trend analysis. The data, combined with sophisticated business intelligence tools, give us more visibility into how our businesses are run and enable better decision-making.

- *We're utilizing more capacity-hogging multi-media data* (e.g., online video, security-related storage, medical imaging). Video is the fastest growing Internet application, as video-based communications extend far beyond traditional media and entertainment companies.

- *We are sharing more data*, which requires real-time data to be accessible to multiple applications and constituents. Collaboration is the new way of doing business, with collaboration occurring not only among a company's employees, but also with partners, suppliers, and customers worldwide.

- *We are maintaining duplicate databases*, for multiple reasons: e.g., to support our business continuity and disaster recovery plans; to maintain high availability; to comply with regulations; to enable multiple applications to use the

data; or even for parallel test or training environments. As a result, a single database may be replicated dozens of times, with each copy maintained separately.

The trends reflect changes to the way we do business—and they're not going to slow down or stop any time soon.

THE HIDDEN COST OF DOING NOTHING

Perhaps you're thinking data growth isn't a problem for your company, at least not right now. Your database manager seems to be taking care of things adequately, and you haven't noticed any unexplained budget upticks.

As long as your database managers find room in the data center to mount additional storage, and as long as you keep approving the costs, you may not be aware of the problem. But that doesn't mean it's not there. In fact, the cost-related impact of the storage problem can easily be hidden.

Continuing to throw hardware at the problem is, at best, a short-term solution. Eventually, you'll run out of data center space. And storage capacity, while declining in price on a per-GB basis, isn't free. With data growing at 20 to 40 percent a year, the cumulative cost adds up quickly.

Consider the true cost of unaddressed data growth. First, of course, is the top-line cost. Your storage budget will reflect storage systems for the primary data plus the multiple backups you maintain. As the storage volume increases, data center costs (e.g., power and cooling), maintenance fees, and labor costs will also increase exponentially. *A rule of thumb is that each 1 GB of data stored represents $1000 in annual costs. Each TB can add $1 million to the budget.*

But the cost to your business goes beyond dollars. In an inefficient storage environment, the sheer volume of data will take its toll on application performance. Response times will slow, as query-related processes are forced to slog through increasing amounts of (largely irrelevant) data. Page loads will be delayed. Backup and recovery will take significantly longer—which may make key applications inaccessible for unacceptably long periods of time. This will lead to productivity losses, as employee, partner, and customer transactions take longer across all applications that access the voluminous database.

At the same time, businesses that don't efficiently manage their data can suffer further losses—to their reputation. Demanding customers are quick to complain about degradation of service (who among us feels forgiving when a customer care rep says, "sorry, the system is so slow today"?). Outages—even during planned backup windows— may receive unwanted media attention. Missed service level agreements may cause not only dissatisfaction, but also bottom-line-impacting payouts to customers.

DEFINING AND ADDRESSING THE STORAGE PROBLEM

So, a tidal wave of data is rapidly approaching your data center door—and it will continue to gain steam with no end in sight. What's a data center manager to do?

It's probably not a good idea to tell the senior leadership team you refuse to accept any more data for storage. First off, that strategy wouldn't work. You can't stop the flood. And second, you'd lose your job. Data is good for the business. And the job of the storage team is not, at the end of the day, simply to find places to store bytes of data. It's to ensure that accurate, up-to-date data is available for constituents when they need it—and to do so in a cost effective, scalable, and easy to manage environment.

Fortunately, it doesn't take much to start to achieve the storage efficiencies that will help you meet your goal. Just follow three key tenets:

1. Reduce the amount of data you're storing.

2. Keep data in the right place.

3. Optimize utilization of your existing storage infrastructure.

Reduce the amount of data you're storing

Thanks to the quirks of digital data, you don't actually have to store every bit. Instead, you can use data reduction technologies that store the minimal amount of bits required to make the data usable to users and applications. Data reduction techniques work because our data is surprisingly predictable and prone to repetition of patterns. By utilizing software that recognizes patterns and eliminates redundancies, you can shrink the total volume. By reducing the stored data, you will require less storage hardware and reduce backup and recovery times.

There are two primary techniques for reducing data: compression and deduplication.

- ***Compression*** compacts the overall footprint of data stored. Compression can be likened to those storage bags that let you reduce a queen-sized comforter to the size of a handkerchief by sucking out the air with a vacuum cleaner. Like the useless air in the storage bag, most data contains some extraneous or unnecessary bits. By removing them via compression, you can reduce the space occupied by your data.

  Compression is a smart solution for data at every storage level. Unfortunately, most commercial compression products introduce delay when users access the compressed data, which makes them best used for less active or archived data. For active data, you need a compression product that allows users to access data without experiencing delay. ***IBM Real-Time Compression Appliances are uniquely designed to compress primary, online data in real time***. Using patented compression algorithms, the IBM appliances enable you to store up to five times as much data, without suffering performance degradation or delay.

- *Deduplication* technologies can eliminate redundancies among stored documents (e.g., multiple copies of the same email attachment) or within documents (e.g., redundant data strings at the block level), by replacing redundant data with pointers. Deduplication is the data equivalent of sorting through your workbench and discovering you have six Phillips head screwdrivers. If you keep your tools reliably and efficiently organized so you can always get your hands on the screwdriver when you need it, you can get rid of five of the screwdrivers and free up space for new tools.

  By clearing away duplicate files and/or strings of data, you can reduce overall storage requirements. Look for dedupe solutions that reduce storage capacity requirements, while maximizing data throughput for backup and restore. *IBM ProtecTIER Deduplication appliances can reduce backup data on disk by a whopping 25:1 ratio.* Furthermore, by supporting speeds up to 1000 MB/second, you can improve backup and restore performance, thus increasing productivity. ProtecTIER Deduplication appliances are available in a range of sizes, including models designed for mid-sized businesses, and are scalable to accommodate future data growth.

Case Study:

Healthcare Provider UPMC Cures its Storage Ills

UPMC, a global non-profit healthcare company, strained its data center capacity when its storage requirements increased by more than 300 percent. Building new space was not a viable option; it would have been prohibitively expensive. Instead, the organization turned to IBM for a solution that would reduce its volume while maintaining a secure storage profile. UPMC deployed an efficient and high performance storage solution consisting of dual 100 TB IBM System Storage ProtecTIER Deduplication Gateways and three IBM System XIV storage units, all managed via Tivoli Storage Manager. The results were staggering. Deduplicated Oracle databases were reduced by a ratio of 24:1; backup time was reduced by 20 percent, and recovery time was reduced by 50 percent.

Keep the data in the right place

In your kitchen, the coffee mugs are probably kept at eye level, the mixer is tucked away, and the turkey platter is stored in the attic. Your storage should follow the same principle. Data that is most commonly accessed should be stored in a way that is easily accessible and stands up to high traffic demands. Older, less commonly accessed data can be moved into lower-performance storage tiers.

Storage, of course, comes in a range of media, and each medium represents a price/performance tradeoff. By using a tiered storage approach, managed by intelligent software, you will ensure that data is in the optimal place, giving you appropriate performance levels at the lowest possible cost. The key is to have the right media in

place, and to build processes to ensure that the data is moved along the continuum. Here are some tips for ensuring your data is in the right place:

- Understand how your data is being used. Rule of thumb: Newer data (between 1 and 7 days old) is accessed most frequently. So, it's appropriate to invest in a higher-cost solution that offers high availability and easy access, like solid state.

- As data ages, it is accessed less frequently (for example, 70 percent of data hasn't been accessed in more than 90 days). This data can be stored less expensively in progressively lower storage tiers, which may be disk-based or tape.

- Look for a storage solution that automates data migration among tiers, so you don't have to rely on technicians to manually activate the move. That makes the solution more scalable and cost-effective. *IBM Easy Tier technology automates the movement of data between costly solid state and less expensive disk-based systems, improving performance and reducing costs.* Easy Tier monitors stored data for activity, automatically moving "hot" data to solid state and less active data to disk media. For additional convenience, Easy Tier optimization technology is integrated into IBM's flagship XIV and Storwize storage systems.

Case Study:

Lego Builds a Model Storage Solution

Facing double-digit annual sales growth, perennially popular toymaker Lego Systems looked to overhaul legacy operations systems that were insufficient to support the award-winning product design and development group. To increase scalability and productivity, the company chose to implement the SAP business suite on an integrated IBM infrastructure consisting of IBM System Storage and System x Technology, running on IBM POWER Servers, managed via Tivoli software. Redundant IBM System Storage DS8700 servers, placed in separate data centers, provide 1.5 TB of production data, as well as the resilience needed to keep operations running. The flexible tiered storage ensures superior data throughput, reduced costs, and a simplified storage environment.

Optimize your existing storage infrastructure

Before you make yet another storage purchase, first take a good look at your current infrastructure. If you haven't incorporated virtualization, you could be running your storage at less than 50 percent capacity. To add insult to injury, these underutilized storage servers are still taking up full floor space and using the full complement of power and cooling resources.

Storage virtualization decouples physical and logical storage. In a virtualized environment, multiple physical storage devices—regardless of manufacturer—are pooled, so that available capacity can be efficiently shared among virtualized storage data. The entire pool is managed via an administrative console as though it is a single highly efficient

storage device. By virtualizing your storage, you gain immediate benefits in cost avoidance (by deferring new hardware purchases); cost reduction (reduced energy costs); faster deployment of new storage; and ease of administrative management.

Start your virtualization effort by selecting a solution that allows you to get more from your existing storage hardware. **The IBM SAN Volume Controller allows you to create and manage a virtualized storage environment using storage that's already on the floor, regardless of manufacturer.** Increasing utilization as much as 30 percent, the SAN Volume Controller maximizes existing investments and can extend the life of some of your hardware.

When it's time to refresh or replace hardware, look for an energy-efficient, high density storage system that incorporates virtualization. IBM offers a range of storage solutions that integrate virtualization capabilities. With built-in virtualization functionality, the **IBM XIV Storage System is designed to operate efficiently at greater than 90 percent utilization**, *providing a low total cost of ownership.* **The Storwize v7000 midrange disk system includes a virtualization engine that extends to the entire storage environment, reducing costs and reclaiming unused storage capacity.**

Case Study:

American Crystal Sugar Spins Up a Sweet Storage Solution

Minnesota-based American Crystal Sugar, like many agricultural companies, is subject to federal regulations limiting production volume. Prohibited from increasing sales volumes, the company focuses on cost-cutting to boost profits. When an aging data center infrastructure stymied plans to deploy productivity-enhancing SAP software, the company turned to IBM and its IBM Premier Partner, Evolving Solutions, for an answer. The company replaced its ragtag collection of aging storage with IBM XIV storage, with 27 TB of capacity. The older equipment gained a new life as backup and recovery in a secondary data center. Through virtualization, the new solution provides more storage capacity in less space, and is scalable to up to 79 TB. In addition, the solution reduces data center costs and enables faster deployment of software and storage. With Tivoli Storage Manager, the company can manage both the primary and secondary data center environments via a single console, thus reducing administrative complexity as well.

PUTTING IT ALL TOGETHER

Given the magnitude of the data onslaught—and the criticality of enabling persistent access to the data—none of these solutions are sufficient individually. You can't afford to continue to implement short term band-aid solutions. Instead, you need to develop a comprehensive storage strategy—one that will maximize efficiency, scale with your business growth, and is flexible enough to accommodate changes in business requirements.

With its extensive portfolio of offers (hardware, software, and services) and its long history of storage leadership, IBM is uniquely positioned to provide the solution needed. IBM's range of storage devices can be integrated into an efficient storage architecture that meets the cost, performance, and scalability requirements of any sized business. For added ease of implementation, the company also offers pre-integrated solutions. Thus, for example, you can deploy IBM Storwize v7000, which already incorporates Easy Tier tiering technology and external virtualization capability; and combine it with ProtecTIER Deduplication for maximum efficiency in a single storage system.

In addition, IBM experts and IBM Premier Partners, certified as storage professionals, are available to provide guidance as you start your project. This includes helping you develop a needs assessment; designing an efficient and scalable storage environment based on your needs and budget; and providing implementation services.

Stratecast
The Last Word

What's your strategy to deal with the rising tide of data? For too many IT leaders, the strategy seems to rely heavily on hope: that we're wrong about the magnitude of the data influx; that your database team can keep pace; that employees won't complain about the slowdown of operations; that senior leadership won't notice the steadily rising costs.

But there's no reason to wait for data Armageddon before taking action. With minimal investment, you can start implementing steps today to bring efficiency to your storage environment. The results can be astounding:

- **Through IBM file compression technology, you can store five times as much data in the same space.**

- **IBM's deduplication solutions reduce backup data by up to 25 times.**

- **IBM tiering solutions can improve data performance by 3 times, while minimizing use of the most costly storage media.**

- **IBM virtualization solutions increase capacity utilization from an average 50 percent to 90 percent—without disrupting performance.**

The clear benefits to creating storage efficiencies include reduced costs and increased operational performance. In addition, ease of administration lightens the management burden, and scalability means faster deployment of storage.

Solutions that reduce operating and capital expense budgets, that are easily deployed and managed, and increase performance company-wide. With that arsenal at hand, there's no excuse to surrender to data Armageddon.

Lynda Stadtmueller

Program Manager –
Business Communication Services
Stratecast (a Division of Frost & Sullivan)
lstadtmueller@stratecast.com

CONTACT US

Beijing

Bengaluru

Bogotá

Buenos Aires

Cape Town

Chennai

Delhi

Dubai

Frankfurt

Kolkata

Kuala Lumpur

London

Manhattan

Melbourne

Mexico City

Milan

Mumbai

Oxford

Palo Alto

Paris

Rockville Centre

San Antonio

São Paulo

Seoul

Shanghai

Singapore

Sydney

Tel Aviv

Tokyo

Toronto

Warsaw

ABOUT STRATECAST

Stratecast assists clients in achieving their strategic and growth objectives by providing critical, objective and accurate strategic insight on the global communications industry. As a division of Frost & Sullivan, Stratecast's strategic consulting and analysis services complement Frost & Sullivan's Market Engineering and Growth Partnership services. Stratecast's product line includes subscription-based recurring analysis programs focused on Business Communication Services (BCS), Consumer Communication Services (CCS), Communications Infrastructure and Convergence (CIC), OSS and BSS Global Competitive Strategies (OSSCS), and our weekly opinion editorial, Stratecast Perspectives and Insight for Executives (SPIE). Stratecast also produces research modules focused on a single research theme or technology area such as Connected Home (CH), MS and Service Delivery Platforms (IMS&SDP), Managed and Professional Services (M&PS), Mobility and Wireless (M&W), and Secure Networking (SN). Custom consulting engagements are available. Contact your Stratecast Account Executive for advice on the best collection of services for your growth needs.

ABOUT FROST & SULLIVAN

Frost & Sullivan, the Growth Partnership Company, enables clients to accelerate growth and achieve best-in-class positions in growth, innovation and leadership. The company's Growth Partnership Service provides the CEO and the CEO's Growth Team with disciplined research and best-practice models to drive the generation, evaluation, and implementation of powerful growth strategies. Frost & Sullivan leverages 50 years of experience in partnering with Global 1000 companies, emerging businesses and the investment community from more than 40 offices on six continents. To join our Growth Partnership, please visit http://www.frost.com.