

The logo for syncsort, featuring the word "syncsort" in a bold, blue, sans-serif font with a horizontal line above and below the text. The background of the top half of the page is a collage of images: a cable-stayed bridge tower and cables, a bridge deck with yellow lighting, and various shades of blue, green, and light blue squares.

syncsort

RETHINK the Economics of Data

Featuring research from

Gartner

RETHINK the Economics of Data

2
Introduction

3
Go Beyond Deduplication with NetApp
Syncsort Integrated Backup

5
Customer Story: Community Health
Center of Snohomish County

7
Customer Story: Bertelsmann Stiftung
Gains Faster Protection at Lower Cost

8
From the Gartner Files:
Best Practices for Addressing the
Broken State of Backup

18
About Syncsort

Gartner's message is clear: Backup is Broken. In the face of massive data growth, disruptive server virtualization technologies, and ever more stringent recovery service level agreements, IT organizations are struggling with decades old technologies that are fundamentally incapable of dealing with the variety of today's data protection needs. Now, more than ever, a new approach is required.

What follows is a thought-provoking research report from Gartner, titled "Best Practices for Addressing the Broken State of Backup." Based on thousands of end user interviews, Gartner offers a roadmap toward rethinking data protection in ways that maximize value and meet business-critical needs.

Syncsort and NetApp believe that Gartner's insights confirm precisely the issues we have seen first-hand in the market. Data protection is simply not being addressed properly even with significant portions of organizational IT budgets dedicated to solving this challenge. Dollars continue to chase the latest technologies, but the result is not a cohesive protection strategy. Rather, it is multiple silos of technology that overprotect some data sets while under protecting others, and often with no direct correlation to business value.

To address these challenges, Syncsort's unique backup and recovery capabilities were combined with the industry leading features of NetApp FAS storage to produce NetApp Syncsort Integrated Backup (NSB). Designed specifically to resolve key business problems, NSB brings together all the latest and most effective protection methods into a single, truly unified solution that leverages best-in-class technology from two long-time industry innovators. The result is an uncommonly efficient data protection solution that goes beyond deduplication and instead solves all your critical data protection needs.

We hope you find the Gartner report interesting, informative and most of all, useful, as you begin to re-think your data protection solution.

Enjoy the report with our compliments,

W. Sean Ford
Chief Marketing Officer
Syncsort, Inc.

RETHINK the Economics of Data is published by Syncsort, Inc.. Editorial supplied by Syncsort, Inc. is independent of Gartner analysis. All Gartner research is © 2011 by Gartner, Inc. All rights reserved. All Gartner materials are used with Gartner's permission. The use or publication of Gartner research does not indicate Gartner's endorsement of Syncsort, Inc.'s products and/or strategies. Reproduction or distribution of this publication in any form without prior written permission is forbidden. The information contained herein has been obtained from sources believed to be reliable. Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Gartner shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The opinions expressed herein are subject to change without notice. Although Gartner research may include a discussion of related legal issues, Gartner does not provide legal advice or services and its research should not be construed or used as such. Gartner is a public company, and its shareholders may include firms and funds that have financial interests in entities covered in Gartner research. Gartner's Board of Directors may include senior managers of these firms or funds. Gartner research is produced independently by its research organization without input or influence from these firms, funds or their managers. For further information on the independence and integrity of Gartner research, see "Guiding Principles on Independence and Objectivity" on its website, http://www.gartner.com/technology/about/ombudsman/omb_guide2.jsp.

Go Beyond Deduplication with NetApp Syncsort Integrated Backup

As you will see in the following report, based upon thousands of end user interactions, Gartner concludes that enterprises, both large and small, are facing enormous challenges in protecting data.

You might ask, "How did we get here?", and "Haven't we been protecting data for decades?"

The answer is "yes we have", and that is a large part of the problem. Most users are still deploying technology that was designed decades ago when a large server hard drive was 100 MB. Meanwhile, data volumes have exploded. On average, the typical organization sees its data double every 18 months. The typical backup environment moves terabytes of data daily and that creates an enormous amount of stress on the backup process: at the server, across the network and at the target, whether disk or tape.

The nearly universal adoption of server virtualization further complicates the problem. While consolidating multiple workloads onto a single physical server drives substantial cost savings, it stresses traditional backup systems that have always relied heavily on excess system resources to do their work. The practical impact is that

“ We’d often get requests from our developer group wanting a SQL restore, by the time we figured out whether or not the restore was even going to work, it could turn into a three-day process. Now, we get it done in an hour. ”

Jason Bourque
Director of Enterprise Systems
Kerzner International

backup limits the number of virtual machines that can be deployed on a given server, which directly impacts ROI.

These problems are not minor, and they won't be solved by tweaking current processes, such as adding more tape capacity or deploying a deduplication target device. As the report notes:

“A common discussion with organizations regarding backup often centers on the need for improved solutions that will not only incrementally improve the overall state of backup, but dramatically improve it. Organizations are voicing the opinion that backup needs to improve a lot, not just a little.”

“ We required a solution that could go beyond deduplication and help us solve a broader set of data protection challenges. NetApp Syncsort Integrated Backup will provide us with a truly integrated solution that we expect will deliver significant value to our business in the form of secure, faster, more reliable and more efficient backup and recovery. ”

Vice President, System Architecture Manager
First Community Bank

The need is for transformational change that encompasses the full, end-to-end data protection process, and that process begins at the most critical point – the application server. With the need to process data 24x7, the long-standing concept of the “backup window” is shrinking to invisibility. Systems are never off-line, and there is never a good time to stress them with the high-impact backups of old that run for countless hours, often failing in the process.

Recovery is the flip-side of the same coin. Downtime is unacceptable, and spending hours – even days – to bring application data back on-line can no longer be tolerated, whether those applications are internal or customer facing.

Overlaying the entire data protection process is management. Users complain of too many systems, too much management time, and too much complexity. Brittle processes and patch-work solutions fail regularly, exacerbating an already tenuous situation.

Transformational change. Fundamental re-design. Eliminating inefficiencies. Bottlenecks and management overhead. This is the new paradigm for data protection.

“In my experience, NetApp Syncsort Integrated Backup is one of the most logical and intuitive products. My staff does not need extensive training or resources to administer it, allowing us to gain efficiencies and spend time on more proactive and strategic initiatives.”

Robert Scott
IT Manager
Community Health Center of Snohomish County

Beyond Deduplication

Data deduplication, while a key component of any solution, by itself will not deliver the necessary value. Users need to move *Beyond Deduplication* to a comprehensive solution that incorporates data reduction and touches all aspects of the data protection architecture.

To address these needs, Syncsort has joined with storage leader NetApp to deliver a simple, integrated and proven data protection solution called *NetApp Syncsort Integrated Backup* (NSB). Re-thinking the data protection process from a comprehensive viewpoint, NSB delivers the fastest, easiest to use and most resource efficient data protection solution for physical and virtual servers running in mixed storage environments.

With NSB, the barriers to successful data protection fall away.

- Backups that took hours to complete now complete in minutes.
- High-impact client processing is eliminated, maximizing resource ROI
- Backups run multiple times a day for better recovery point protection
- Processing power is freed up for application use, especially critical for virtual machine environments
- Instant data recovery limits downtime by 98%
- A simplified architecture and ease-of-use features cut management costs by 50%

A number of game-changing data protection technologies have been available for some time, but never in one place and with the operational ease-of-use and cost efficiencies delivered by NSB. Only by joining together the best-in-class technologies of two leading industry innovators could a single solution deliver the power and breadth of NSB, while at the same time keeping to the core solution values of *simple, integrated and proven*.

For more information on what NSB can do for your organization, please contact your local NetApp or Syncsort sales person or partner, or visit us at www.syncsort.com.

Source: Syncsort

Customer Story: Community Health Center of Snohomish County

Industry

- Non-profit provider of primary medical, dental and pharmacy services.
- Founded in 1983 to serve the health needs of county residents who face barriers to health care.

Organization Need

- An enterprise data protection solution for their virtual environment
- Need for P2V migrations

Solution

- NetApp Syncsort Integrated Backup (NSB)

Business Value

- Minimized impact to operations staff
- Ability to rollout applications faster
- Business continuity.

The following story appeared on March 8, 2011 in SearchDataBackup.com

["Community Health Center Changes Data Protection Strategy with Syncsort"](#), by Dave Raffo, Senior News Director

Customer Story

A proliferation of virtual machines and digitized patient data prompted Community Health Center (CHC) of Snohomish County, WA, to change its storage management and data protection strategy. The medical center network switched to NetApp FAS storage last year along with NetApp Syncsort Integrated Backup (NSB) to enable disaster recovery, speed backups and the provisioning of storage.

The Snohomish CHC provides medical, dental and pharmacy services to lower income residents without healthcare

“We required an integrated, easy to use solution that could help us achieve fast, reliable backups to more effectively support the business. NetApp Syncsort Integrated Backup offers a compelling value proposition and makes perfect sense for a heterogeneous environment like ours.”

Robert Scott
IT Manager
Community Health Center of Snohomish County

insurance. IT manager Bob Scott said data growth due to electronic medical record (EMR) systems and large-scale virtual server implementation played havoc with his backup windows and the ability to quickly provision storage. He said using NetApp storage with NSB to back up to disk drastically reduced his backup windows and the time it takes to create a virtual testing environment.

Scott said CHC first decided to switch storage area networks (SANs) around the middle of 2010 because it found his EMC Corp. Clariion system difficult to manage for his six-person staff. After he decided to move to the NetApp FAS3100 platform, he discovered NSB. NSB is developed by Syncsort and NetApp and integrates Syncsort BEX backup software with NetApp's snapshot technology.

CHC has more than 80 virtual servers – mostly VMware – and a Microsoft cluster for his EMR. NSB lets him mirror data on two SANs to provide the ability to failover with no loss of the critical EMR application.

Before switching to NSB, CHC did its daily data backups to tape with Symantec Corp. Backup Exec. Now it uses Syncsort to handle daily backups to disk, including virtual server backups.

“We had outgrown our SAN and were looking for something easier to manage,” Scott said. “We were getting bogged down in backup. We had no service windows because our backups were taking so long. I needed a better solution for backup.”

NSB includes Syncsort's BEX backup software for physical and virtual machines, and is integrated with NetApp snapshot technology. It uses NetApp's SnapVault for bare-metal recovery and SnapMirror to replicate snapshots for disaster recovery.

Now CHC backs up to a 20 TB partition on the FAS3160 with a 20 TB mirror on a FAS3140. Scott said he still has tape for monthly backups because of long-term retention compliance requirements. He said moving to disk cut his nightly backup window by five hours – giving his team more time for maintenance.

“We went from running backups all night from 9 p.m. to 5:30 a.m., and that's if they ran correctly,” he said. “Now we start at midnight and we're done at 3:30 a.m. That gives us a nightly service window from 8 p.m. to midnight that we didn't have before.”

We were backing up every night except for a two-hour window on Sundays.”

Scott said he takes advantage of NSB’s Instant Virtualization to feature create a virtual testing environment. Instant Virtualization lets him convert any backup image into a VMware virtual machine by using a NetApp FlexClone volume instead of through the VMware console. He said the process takes minutes, and uses no additional storage. Because it runs off the backup storage, Instant Virtualization doesn’t impact the production environment.

“The ability to test within the Syncsort application utilizing the NetApp SAN was huge for me,” he said. “We didn’t have the funding to create a sandbox test environment. The Instant Virtualization

within the backup environment gives us a test environment. I can snap a production database, test it, and if it breaks we can blow it away. If it works, we copy it over to production.”

Being able to set up a test environment quickly comes in handy at CHC. Scott said at any given time he might have 60 or so testing projects to support, and many come to him with little or no advance notice. He said it took up to two weeks to re-stripe a LUN on his Clariion storage but no re-striping is required in his current setup because he can make data-in-place point-in-time copies of LUNs with snapshots.

NSB uses Syncsort’s Zero Impact Block Level Backup on the client side to only transfer

blocks of data that have changed since the last backup and takes advantage of NetApp’s target-side data deduplication.

“We dedupe in each aggregate [multiple RAID group] across the board before it hits Syncsort backups, so we’re not backing up more information than we need to,” he said.

He said he came across Syncsort last July when looking for a disk backup product around the time he was switching SANs. “I didn’t believe some of their claims for restores, especially with large Exchange mailboxes and servers,” he said. “I made them come down and prove it to me, and they did. Then I found out about the integration with NetApp. The fact they worked so well together was icing on the cake.”

Source: Syncsort

Customer Story: Bertelsmann Stiftung Gains Faster Protection at Lower Cost

Industry

- Public Sector
- Based primarily in Europe
- The Bertelsmann Stiftung foundation employs about 300 employees.

Organization Needs

- Ability to handle rapid data growth
- Faster recovery in case of disaster
- Optimization of RPO and RTO
- Cost savings

Solution

- NetApp Syncsort Integrated Backup (NSB)

Business Value

- Faster, more efficient disaster recovery solution assures rapid restore and minimal business interruption
- Solution provides a base for higher Service Level Agreements (SLAs)
- Long-term protection needs satisfied, even for increasing data volumes
- Low administration expense
- Expanded reporting options

“With Syncsort and NetApp, we can backup and restore data faster and better than was possible in the past.”

Klaus Brinkmann
IT Systems Coordinator
Bertelsmann Stiftung

Customer Story

Interested in identifying major societal challenges and developing possible solutions, the Bertelsmann Stiftung foundation has an abundance of information available to them. Data protection and recovery play a vital role in their operations. Due to the recent influx of data, backups had become extremely time-consuming and, with the solution they had in place, could only be executed once per work day. Recoveries also were taking an excessive amount of time.

Concerned about data growth, the Bertelsmann Stiftung IT team began searching for an alternative data protection solution, and they quickly found one in Syncsort and NetApp. NSB operates at the block-level, minimizing the time and disk space needed for data protection as well as accelerating the recovery process. “This combination met our requirements exactly,” recalls Klaus Brinkmann, IT Systems Coordinator of the Bertelsmann Stiftung. “We interviewed other companies who were already using the Syncsort solution and received outstanding references.”

“As long as the backups continue to function flawlessly, one assumes that the recovery process is also working,” observes Klaus Brinkmann. “However, we wanted to know for sure.” Intensive testing was performed for two long months. The result was conclusive: NSB was the right way to go.

What became immediately apparent was that NSB significantly reduced recovery time, which reduced downtime cost and delivered measurable value.

Christoph Riedel, managing director of Trading. Point GmbH, the organization responsible for setting up Bertelsmann Stiftung’s backup environment notes that NSB “has enabled the Bertelsmann Stiftung foundation to recover lost data in a timely manner.” Open Systems SnapVault (OSSV) technology, supported by NSB, played a critical role in fast, efficient data protection and recovery. Using this technology, multiple data backups can be performed daily as snapshots. OSSV reduces the time necessary for data backup and recovery by up to 95 percent.

Source: Syncsort

Best Practices for Addressing the Broken State of Backup

This research details the challenges associated with the current state of backup and recovery, and presents the top best practices for improving the backup process.

Key Findings

- Backup issues such as cost (acquisition, but especially maintenance), product capabilities and management complexity are among the top pain points organizations complain about.
- Just as there is tiered storage that provides a variety of cost and capabilities for primary and secondary storage capacity, the concept of “tiered recovery,” which provides differentiated levels of backup/recovery services according to business value, is gaining traction.
- The concept of “unified recovery management” is evolving, whereby numerous backup techniques, and potentially backup solutions, are centrally managed from one console.

Recommendations

- Develop an updated and comprehensive backup/recovery plan for all the organization’s current data, and build in assumptions on data growth and architecture changes that are anticipated in the next three years.
- Look to deploy, or more fully deploy, recent proven enhancements in available backup products, such as incremental forever, synthetic or virtual full processing, deduplication, server virtualization improvements, and snapshot and replication integration.
- Build the concept of tiered recovery into your availability strategy, establishing multilayered SLAs that appropriately match recover techniques to business requirements.

- Implement a unified recovery management solution where ideally all, but at least many, of the backup techniques and approaches are managed from a common and centralized point.
- Do not use the backup system for long-term data storage or archiving; instead, implement separate archive systems for long-term data retention.
- Perform data recovery testing at least once a year on a subset of your data to ensure that your backup strategy can effectively meet the stated protection SLAs.

STRATEGIC PLANNING ASSUMPTION(S)

By 2013, at least 20% of organizations will have changed their primary backup vendor due to frustration over cost, complexity and/or capability, up from the typical single-digit percentage shifts today.

By 2013, more than 50% of midsize organizations and more than 75% of large enterprises will implement tiered recovery architectures.

ANALYSIS

1.0 Introduction

Backup and recovery is one of the oldest and most frequently performed operations in the data center. The need for recovery solutions dates to the advent of computers themselves, and backup has been an established market for many decades. Gartner end-user inquiry call volume regarding backup continues to be high every year, and has been rising at about 20% each year for the past three years. Organizations worldwide are seeking ways to easily, quickly and cost-effectively ensure that their data is appropriately protected.

Despite the long timeline associated with backup, the practice has seen a number of changes and challenges in recent years.

Although new backup techniques and technologies have come into the market, many organizations are still struggling to partially embrace, much less fully deploy, newer backup capabilities. At the same time, a common discussion with organizations regarding backup often centers on the need for improved solutions that will not only incrementally improve the overall state of backup, but dramatically improve it. Organizations are voicing the opinion that backup needs to improve a lot, not just a little. This is leading to an increased willingness to switch backup vendors. The implication with the rising frustration with backup is that data protection approaches of the past may no longer suffice in meeting current, much less future, recovery requirements.

2.0 The Broken State of Backup

For many organizations, backup has become an increasingly daunting and brittle task fraught with significant challenges. In this section, we present a closer inspection of the greatest contributing factors to frustration in the backup/recovery process.

2.1 Amount of Data

The industry has talked about the challenges associated with the proliferation of data for years. While it is generally accepted that the amount of data is growing fast for most organizations, a closer look at data growth reveals why this is a major issue for backup.

Structured data, such as databases, and even semistructured data, like e-mail, usually have an administrator charged with managing the data and the applications that go with it. Growth is an issue in this case, but even more challenging is the typically faster growth rate for unstructured data, such as office productivity files (e.g., word processing, spreadsheets and presentations), which do not have a central administrator. The lack of a central owner and administrator, coupled with the high rate of redundancy and accidental corruption and deletion of unstructured data, make appropriately protecting these files even more challenging.

Data growth is not only about the number of objects (files, data base, e-mail, SharePoint, virtual machines [VMs], etc.) increasing. The size of each of these objects is also on the rise, as e-mail becomes an unintended content repository, databases now house embedded files, and spreadsheets and word processing documents contain rich media files. Additionally, some organizations would like to keep multiple versions of objects in order to revert to previous instances, further adding to the amount of data to back up and retain.

In total, data growth is occurring from increases in:

- The number of objects
- The size of objects
- The number of versions
- The desire to retain these objects (files, e-mail, etc.) for longer periods

Each of these adds to the amount of data to back up, especially if an archiving solution is not implemented to deal with aged files, and/or data reduction techniques (such as data deduplication) are not deployed.

2.2 More-Stringent SLAs

SLAs are increasing in most organizations and are becoming more stringent, as well as more common, in terms of an increased number of companies that have backup service levels. The desire today is for zero downtime and to recover data faster than before. For some companies, the notion of a 24-hour, or longer, restore process is no longer acceptable for any data, not even personal files.

In the last five years, a desire for improved recovery time objectives (RTOs) and recovery point objectives (RPOs) have been well-documented. RTO is how long it takes to recover from a given failure scenario, and RPO indicates how far back in time the

restored data will be from. Perhaps the most accurate way to view these is that, from a business perspective, RTO indicates how long you're going to be without the data, and RPO is a measure of how much of the most recent data will be lost. Naturally, organizations are striving to make improvements in both measurements for as much of their data as possible.

As is to be expected, most lines of business and end users would like to achieve the highest level of service, which means the most stringent RTOs and RPOs. In a perfect world, the premier recovery methodology would be applied for all data. In reality, constraints such as budget, time and storage personnel translate to the need to relegate some data to a lower quality of service, resulting in longer recovery and more potential for lost data. The real issue is that the business or end user may not be fully aware of the true level of service, often believing they are better protected than they are.

The desire for better protection and availability results in a demand for improved backup capabilities.

2.3 Antiquated Backup Model

For several decades, the backup process has remained largely unchanged, using the same techniques and methodologies that were used over one-quarter of a century ago. Classically, the backup model was that once a week a full backup would be performed (often over the weekend), making copies of all data. Then, once a day (typically late in the evening, when the data would presumably not be accessed), an incremental backup process would make a copy of new and modified files, databases, etc. Some organizations were concerned about the restore process taking too long to recover a full backup and potentially having to apply several nightly incremental backups to get to the most recent data, so to avoid this a nightly full backup might have been performed on critical databases or e-mail systems.

This backup model represented a brute-force approach, where quite a bit of server processing, network transmission and storage of redundant data took place. While this may not have been a significant issue decades ago, the growth of data, proliferation of new applications and the need to access data around the clock eventually exposed the traditional backup methodology as broken. The legacy backup approach simply does not scale well, and cannot handle the increased demands of protecting more data in less time, ideally with faster recovery.

Scale issues often reported with traditional backup have two types of challenges:

- The first issue is the "very large object problem." This happens when a company has a very large database to back up. An example of this is when a large database table space, perhaps several hundred terabytes in size, must be backed up regularly, but the full backup may take over 24 hours to complete. This means a backup may not complete before the next backup is scheduled to take place. Using legacy methods of backup can take a long time, perhaps too long for some organizations to obtain a full backup.
- The second issue is the opposite concern. Instead of a smaller number of very large objects, an organization may have millions, or even billions, of relatively small files, perhaps only a few kilobytes in size. The problem is that while the amount of data is very small, and can thus be quickly backed up, the extremely large number of files takes a long time to process. This is because most backup applications need to traverse the file system, scanning for new and modified data, and then create an entry for the object in the backup repository/database. These millions of database entries can take up more space than the actual object, and can slow the backup process. The scanning process consumes

significant processor cycles and a long time to complete. Gartner has heard nightmare accounts where the time required to interrogate the file system could be hours, but that the time spent backing up new and modified data could be a matter of minutes. The end result for this issue is the same as the first – the backup takes too long to complete.

A further concern with most traditional backup approaches is that backup agents for file systems, databases, e-mail, ERP systems, etc., had to be deployed on each physical machine in order to obtain a backup. The rise of server virtualization initially made this even more daunting, as even more agents needed to be installed and managed. In larger enterprises, the challenge of keeping track of and updating backup agents was made more problematic because the storage team might have trouble getting server administrators to allow them to update and maintain the backup agents, meaning that down-level backup agents had to be endured for long periods of time.

Snapshot and data deduplication techniques have recently proved useful in addressing backup scale challenges, and are increasingly being deployed as a means of improving the level of efficiency in the backup process. The penetration rate for both of these technologies, while on the rise, is estimated by Gartner to be less than 20% for large enterprises.

2.4 Top Backup/Restore Frustrations

Each year, Gartner analysts collectively handle close to 1,000 inquiries on backup, and host several conferences in multiple countries where backup issues are discussed. The following are the top three frustrations that end users have expressed over the last two years.

2.4.1 Cost

First and foremost, organizations complain about the cost of backup solutions. The upfront acquisition cost is always a concern, but the larger issue for many often is the annual maintenance fee associated with backup software. As more applications have been deployed and more machines (physical and virtual) have been installed, the costs of backup have risen. Often coupled with this is the desire to expand backup to include more servers than in the past, to add protection to remote offices, and sometimes to provide new backup services for desktop and laptop users as well. All combined, these things can result in a larger backup bill, and a much larger annual maintenance and service cost.

2.4.2 Capability

Cost may be the most frequently cited backup frustration, but it is often aggravated by the second most common complaint, which is the capability of the backup solution. Organizations become more exasperated by the cost of their backup solution if they feel that the product is not adequately meeting their recovery needs.

From end-user inquiry, conference polling and worldwide surveys, the top five backup capability complaints in priority for backup solutions are:

- Difficulty meeting the backup window
- The need for point solutions in order to fully protect all systems or data (e.g., server-virtualization-specific tools, unique database or SharePoint protection products, and different solutions to back up laptops)
- Needing to troubleshoot and restart failed backup jobs

- Not being able to restore data fast enough to meet SLAs
- Lack of easy and/or complete reporting to know and prove data is protected

2.4.3 Complexity

Organizations also complain about the complexity of their backup solutions. The feeling is that too much time, expertise and effort is spent keeping the current recovery system(s) afloat. Organizations would like backup to be a process that requires far less supervision and administrative attention, and for higher backup and restore success rates to be easier to achieve.

Some complexity issues have been cited earlier in this research, but common specific examples include:

- Agents:
 - Having to deploy and manage code on each physical, and often each virtual, machine is problematic for large organizations.
 - In some cases, multiple agents need to be deployed on one system to protect the file system, application and network-attached storage (NAS) data.
- VMs:
 - The most common method for protecting VMs is still to put a traditional backup agent in each.
 - While advancements have recently been offered by some hypervisor vendors, a lot of confusion remains in the industry about best approaches, what new capabilities are facilitated by each hypervisor provider, and what level of hypervisor support and VM backup capabilities are supported by a specific backup vendor.

- Cascaded applications:
 - Applications that run on top of other applications can present additional backup concerns and complexity. An example is SAP running on top of an Oracle or SQL Server database, which could require expertise and special handling for each of these components.

- Organizations complain of the lack of powerful, easy-to-use solutions for these environments.

- SharePoint:
 - Emerging applications like SharePoint may not have a current backup plan or best practices associated with them, since they are so new.

- Microsoft has offered a new version of SharePoint, which requires the use of different interfaces and best practices to properly protect it.

- Further adding to the challenge is that there are many ways to implement a large SharePoint instance, whereby the many application components could be installed on many different servers.

- Backup beyond the data center:
 - Protection for remote offices, desktops and laptops is on the rise, but brings with it new demands and concerns.
 - A potentially new class of less-sophisticated and geographically dispersed users with mobile devices

that can more easily be lost, stolen or damaged carries unique data recovery and service organization challenges.

- Many data centers view this new backup task as the tip of the iceberg, as client virtualization, or hosted virtual desktops (HVDs), bring new and expanded recovery demands.

3.0 Best Practices for Improving Backup

Gartner is often asked how organizations can improve their backup and recovery practices. Backup may not be glamorous, but it is an activity that nearly every organization is striving to optimize – typically trying to contain costs, while protecting more data and offering faster backup and restore times. In this section, best practices for improving backup are presented.

3.1 Fully Explore Current Capabilities

While backup/recovery solutions certainly have a great deal of room for improvement, it is also the case that most organizations have yet to fully embrace and deploy data protection approaches and techniques that have been available for years. Such concepts as incremental forever, synthetic backups and virtual full backups are now offered by several vendors, and are robust enough for production implementations. The expanded use of disks for backup, while a marketing slogan for some time, is now becoming the norm for the industry (see “Poll Shows Disk-Based Backup on the Rise, With a Few Surprises”). Client- and target-side data deduplication solutions are offered by many providers. In fact, some vendors offer multiple kinds of deduplication, and are on second- and third-generation solutions. The ability to do image-based snapshots with individual,

file-level cataloging and restoration has been possible for quite some time as well. Many vendors have delivered significant improvement in server virtualization and SharePoint recovery in their most recent product releases.

Many vendors have been busy developing code internally, and a number of acquisitions in the last four years have led to expanded features. While acquisitions bring the potential for new capabilities, they can add to integration and poor ease-of-use side effects. Nonetheless, one company acquiring another is a strong signal of the future direction of the acquirer’s strategy, and an indication of where the industry sees the market headed, and is useful in understanding what functions are becoming important.

A partial list of major backup and recovery acquisitions since 2006 includes:

- Atempo
 - Storactive for continuous data protection (CDP) and laptop backup (2006)
- BakBone Software
 - Asempra for CDP and replication (2009)
 - Alvarri for deduplication (2008)
 - Constant Data for CDP and replication (2005)
- CA Technologies
 - XOssoft for CDP and replication (2006)

- EMC*
 - Data Domain for target-side deduplication (2009)
 - WysDM for backup reporting (2008)
 - Berkeley Data Systems for Mozy online backup services (2007)
 - Indigo Stone for bare-metal recovery (2007)
 - Avamar for client-side deduplication (2006)
 - Kashya for CDP and heterogeneous replication (2006)
- IBM
 - Diligent Technologies for target-side deduplication (2007)
 - FilesX for CDP and replication (2007)
 - Arsenal Digital for online backup service (2007)
- Quest Software
 - Vizioncore for server virtualization backup (2008)
- Symantec*
 - Revivio for CDP and heterogeneous replication (2006)

*While further back in time, it is worth noting two additional major backup/recovery acquisitions: Veritas by Symantec in 2005 and Legato Systems by EMC in 2003.

New capabilities take time, often three to five years or more, to gain widespread adoption, as most organizations are risk-averse. However, the delay in implementation is

also attributable to companies not knowing all the available options and capabilities their current backup vendor, much less all vendors in the industry, currently offers.

Action Item: Before making plans to jettison your current backup product, make sure you have been fully updated on all the features that have been delivered in the last three years to ensure that you are fully leveraging your investment.

3.2 Implement Archiving

The vast majority of the backup methodologies in production use a model whereby frequent complete full backups of all data are taken. While it is often standard procedure to configure the application for nightly incremental backups for six consecutive days and then once a week (usually on a weekend day) to take a full backup, many organizations opt for a full, nightly backup for e-mail and critical databases to minimize the amount of restore processing that needs to occur. While this approach has worked well for most companies for several decades, many now find they cannot contain all the backup activity in the available time (the backup window). While newer backup approaches can help address this, most backup applications still rely on the “full-plus” incremental concept.

In this case, removing files, e-mail and application data from primary storage can drastically reduce the overall amount of backup processing required during each full backup. Sometimes this is referred to as reducing the “working store.” Organizations might ideally perform a garbage collection process, perhaps using a storage resource management (SRM) tool, but most often using data identification tools in the archive solution itself to identify candidates for archive, or to outright delete unneeded and duplicate data. Often, the storage team is

concerned with deleting data, and they might not have the SRM tools at their disposal to investigate which data is not used or could be deleted. This makes implementing an archive solution that moves data to lower-cost storage devices, often by using a redirection mechanism, a very beneficial method for quickly and effectively reducing the backup window. Additional backup benefits of implementing archiving include faster restore times for a complete recovery of all backed-up data (since there is less data to bring back), as well as reducing the cost of storing and potentially transporting the backup media off-site.

An added benefit of archiving is that backup retention periods are frequently decreased, since archive is used for long-term retention, and backup is used for operational recovery. This results in a much lower exposure for e-discovery during a litigation activity, and can help contain the high cost of scanning backup tapes for required legal materials.

Action Item: Implementing an archiving procedure can be the most cost-effective way to improve backup and restore time, and a means of reducing primary and backup storage costs.

3.3 Consider New Backup Pricing Models

3.3.1 Backup Cost History

Backup products have traditionally been priced on a per-server basis, often with many add-on costs for features and line items, with some newer features commanding a premium price. Over time, most vendors have collapsed the number of charged items into their base product, or added features into an expanded/extended/enterprise offering of the product. This practice has been going on for over 20 years. An older example of this is how the number of tape library slots used to be charged in groups, requiring organizations to license several groups or

tiers of tape library slots in order to deploy a large tape solution. Eventually, most backup vendors gave way to dropping all unique pricing for tape libraries and the number of slots in the library. Over time, charging for software compression, encryption, basic disk support, base reporting, etc., has typically been folded into the product.

Action Item: While there can be a risk in upgrading to new backup software versions, if you are under a maintenance agreement, find out whether you are entitled to free upgrades, and consider newer versions when the new release bundles previously charged-for features.

3.3.2 Backup Costs Today

In addition to capabilities being completely bundled into a product, the premium that is charged for features often decreases over time. With this in mind, we should expect current deduplication up-charges to decline, perhaps as a result of competitive pressures from vendors like CA Technologies (ARCserve) and IBM (Tivoli Storage Manager [TSM]), which offer software-based deduplication at no additional charge, causing competitors to waive their additional licensing fees.

While feature collapse and new feature price erosion can sometimes help lower costs, the top complaint about backup solutions is still cost: initial product acquisition costs and ongoing maintenance fees in particular. For larger enterprises, the maintenance costs of the typical three-year backup software purchase agreement are the greatest concern, since they represent a future spending commitment.

Cost concerns are even driving some organizations to consider cloud backup services, which promise a lower total cost of ownership (TCO) for backup.

Action Item: When negotiating with vendors, first understand which features are additionally charged for and what capabilities are included at no additional expense in all products that are on your shortlist, to ensure an accurate TCO and for use in pricing negotiations.

3.3.3 Maintenance Becomes a Factor

Some vendors have been able to prop up their overall revenue partly due to a large installed base from which they continue to derive substantial maintenance revenue. However, Symantec in 2007 and 2009, and IBM in 2009, have received significant end-user backlash regarding perceived increases in the overall cost and method by which maintenance is calculated. This has driven some companies to consider alternative solutions, and has become an attack point by competing vendors. In some cases, swap-out programs are in place whereby competitive solutions are brought in only on the cost of maintenance for the products they are replacing.

Action Item: If you are considering switching backup applications, note that some vendors sometimes will agree to only charging for the maintenance of the new application, waiving the initial acquisition cost.

3.3.4 Vendors Change Licensing Terms and Models

By 2009, most backup vendors introduced pricing updates to reduce the cost of their software in a virtual server environment. Prior to this, many vendors charged the same amount, typically for a much smaller VM as for a physical server.

The most drastic recent pricing change was the introduction of capacity-based licensing. Whereas traditional pricing was based on the number of servers (along with specific server attributes, such as the number of processors

or types of OSs), capacity-based pricing usually charges independently of the type of data and its origin, and is calculated on places a cost per unit of storage (usually per terabyte for enterprise solutions) basis. CA Technologies introduced this in 2009, IBM and Symantec in 1H10, and CommVault is making this available worldwide in 2H10. The vendors can differ as to where the capacity is measured – on the “front end” for the data being backed up, or the “back end,” measuring the amount of data that the backup solution generates.

Action Items: If server virtualization is deployed, make sure your organization gets the benefit of recent pricing changes, like partial server or partial CPU charges. When evaluating new backup solutions or investigating extending a maintenance agreement with your backup vendor, ask the vendor if there are new pricing and packaging plans that are available, such as capacity-based licensing, newer product versions with a collapsed parts list that include additional features in the base product, and bundles that combine many features and/or products at a lower overall cost.

3.4 Implement Data Reduction

Deduplication has become “table stakes” for vendors in the backup/recovery market. The value of data reduction technologies, such as deduplication, cannot be understated. Deduplication materially changes the economics of disk-based backup and recovery approaches by reducing data, resulting in significantly lower disk requirements and cost.

In May 2007, Gartner said that deduplication was a transformational technology with the potential for significant cost savings and expanded quality of service capabilities (see “Data Deduplication Is Poised to Transform Backup and Recovery”). We reiterate this

assessment, and frequently advise clients to investigate deduplication technologies for use in addressing current and anticipated storage challenges. Gartner believes that data reduction, such as compression and deduplication, is a “must have” capability for backup solutions.

The benefits of deduplication are in resource savings. Potential savings can occur on many levels. The primary benefit is in substantially decreasing the amount of disk space required to store a given amount of data. Vendor claims range from capacity savings of 20:1 to 400:1. Gartner clients with deduplication technology in production report data reductions ranging from 8:1 to almost 300:1. Actual ratios vary depending on the amount of data redundancy, the data change rate and the backup methodology (for example, full, full plus incremental, full plus differential or incremental forever). The more often full backups are made, the higher the deduplication ratio.

Depending on the deduplication implementation, there can be bandwidth savings in the amount of data transferred over the network. Other favorable side effects of deduplication include decreased power, cooling and the physical size for storage devices as a result of using less physical capacity and reductions in acquisition costs.

Another key advantage of deduplication is the ability to improve SLAs for recovery. Backup to disk can improve backup windows and enable production resources to return to normal operations faster, minimizing the impact of backup on production activities. However, some organizations have not incorporated disks into their backup environments, and many companies have only deployed a small amount of disks if they are using a disk-based approach. Those that have deployed disks often use it as a staging area or cache, which is soon cleared by transferring data to tape to make way for newer backups.

Deduplication can significantly change the economics of the amount of disk space needed for storing a given period of backup data, enabling users to move from days/weeks to weeks/months of online retention and, in some cases, a tapeless environment. As a result, deduplication can also positively impact restore times, because the longer the ability to house data on disk, the greater the likelihood that a restore request can be satisfied from a disk copy versus a slower access medium, such as tape or optical. There has been concern about the data re-inflation process (also referred to as “rehydration”) adding restore time, but as processing power increases, the deduplication solutions become faster, and comparing current deduplication implementation to typical legacy tape solutions yields a positive improvement in restore time.

Deduplication could enable organizations of all sizes to consider or expand their use of disks as a backup target. With improved disk-based backup comes the ability to restore from disks, improving recovery times.

Deduplication has generated an increasing number of end-user inquiries since 2006. Gartner conference session polling, kiosk polling at conferences and global surveys continue to show an increase each year in interest in and deployment of deduplication. (These recent results support an earlier Strategic Planning Assumption in “Predicts 2008: Emerging Technologies Make Storage Hardware and Software More Effective,” which states that, by 2012, deduplication will be applied to 75% of backups.) Gartner expects increasing industry buzz and end-user interest in deduplication to continue, and to accelerate beyond 2010. The interest in data reduction technologies are well-founded, as they bring many benefits to the recovery process.

Action Items: When evaluating backup software or disk-based hardware solutions, data reduction (such as compression and data deduplication) should be considered a must-have feature, and should be an essential part of assessment criteria. Many backup vendors have released new or expanded deduplication features, so it is important to understand the latest capabilities so as not to be incorrectly swayed by a vendor positioning the competitive capabilities of other solutions.

3.5 Implement Tiered Recovery

In a perfect world, IT would have all the resources required to implement the optimal level of capability and service for all applications, users and data. Unfortunately, this is rarely the case, as constraints such as funding, storage administrator time and the availability of one product to handle all backup requirements are often issues. The need to balance the budget against resources, and to manage availability trade-offs in a business-appropriate manner, are leading to a new backup and recovery approach.

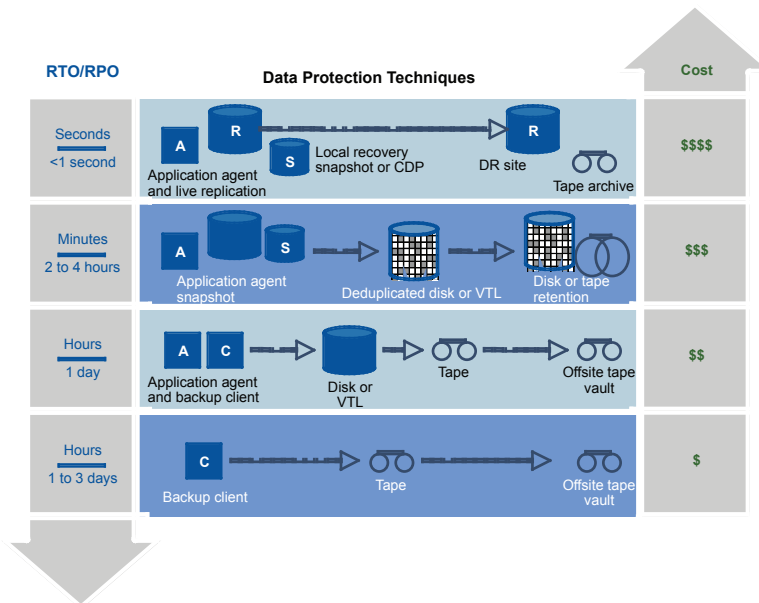
Over the last five years, Gartner worldwide surveys, conference polls and end-user inquiries clearly show that backup is slowly morphing from tape-only to more-disk-oriented activity; however, in the past two years, the industry has begun to embrace and deploy recovery enhancements. Not surprisingly, disk is increasingly being incorporated into the backup process, and in greater amounts. While, directionally, we see disk usage, data deduplication, replication for electronically vaulting off-site copies of data and snapshots for taking more-frequent copies of data all on the rise, the same tools, technologies and backup techniques from decades ago are also typically implemented. This expanded menu of options, techniques, cost structures and service levels has changed the way organizations deliver backup services.

In the past, backup was very much a “one size fits all” endeavor. Often, the only major backup decisions being made were whether something should be backed up and, if so, how long to retain it. The classical approach was a once-a-day copy of the data, often in the midnight hours, which was then written to physical tape. Fast-forward to today, where multiple technologies are often used in conjunction with one another and are merging to provide additional capabilities. New techniques lead to an expanded menu of choices, and one current or emerging recovery technology does not always win out over another. Rather, administrators will have more flexibility, including differentiated levels of cost and service, in providing recovery solutions. Just as the concept of tiered storage provides a variety of cost and performance levels for storage capacity, now tiered recovery provides differentiated levels of backup and recovery services. Unlike tiered storage, the tiered recovery model may be additive, with an organization using multiple techniques together to achieve the needed overall level of data availability and recovery characteristics, and to ensure that business risk and continuity requirements are met.

One example of the tiered recovery concept is an organization that protects its vital databases by performing nightly backups to a virtual tape library (VTL), which are deduplicated and replicated to another company location where they are written out to physical tape and vaulted. This same company also takes snapshots throughout the day at two-hour intervals to further guard against outages, and retains the last 10 snapshots before recycling the space for newer snapshots.

To effectively implement tiered recovery, an organization should conduct a business impact assessment (BIA) to categorize the

FIGURE 1
Tiered Recovery Example



Source: Gartner (August 2010)

criticality of the IT services. Any recovery architecture must start by gaining an understanding of the IT services supporting business processes and their associated service levels. The reason for this is that service levels affect the capabilities, cost, architecture and complexity of the backup solution. This knowledge is typically acquired within a business continuity management (BCM) program – through a BIA. While, historically, this was only performed for disaster recovery systems, Gartner recommends performing a BIA for backup data as well, in order to contain cost and deliver the most appropriate service levels. Most organizations specify three to five tiers of criticality, with the highest tier having

the most stringent service levels. Every organization is different, however. This issue is business-dependent, based on how IT services are utilized in the business processes, and on the overall impact of a disruption to IT services and the availability of data. Tier 1 data has the most stringent RPO and RTO requirements, and the cost of the backup solution is the highest (see Figure 1). Not all applications and data in a critical business process would be grouped in Tier 1; rather, only those deemed most critical or with the most downtime effect (typically around 30% of a portfolio reside in Tier 1; as costs improve, more applications and data can be moved to a higher tier for improved service).

Action Items: Implement tiered recovery to optimize the balance between cost and recovery service levels. Conduct a BIA, and review it annually to determine the criticality of your business systems and their data. Implement tiered recovery by using the BIA results and devising three to five tier categories, and associate recovery service levels to each tier (that is, RTO, RPO, retention, off-site copies, etc.).

3.6 Implement Unified Recovery Management

New and less-expensive disk options make the use of disks for faster recovery a more viable option than backup to tape. Enterprise backup vendors have responded with better disk support (such as disk-to-disk, deduplication and replication), and virtual

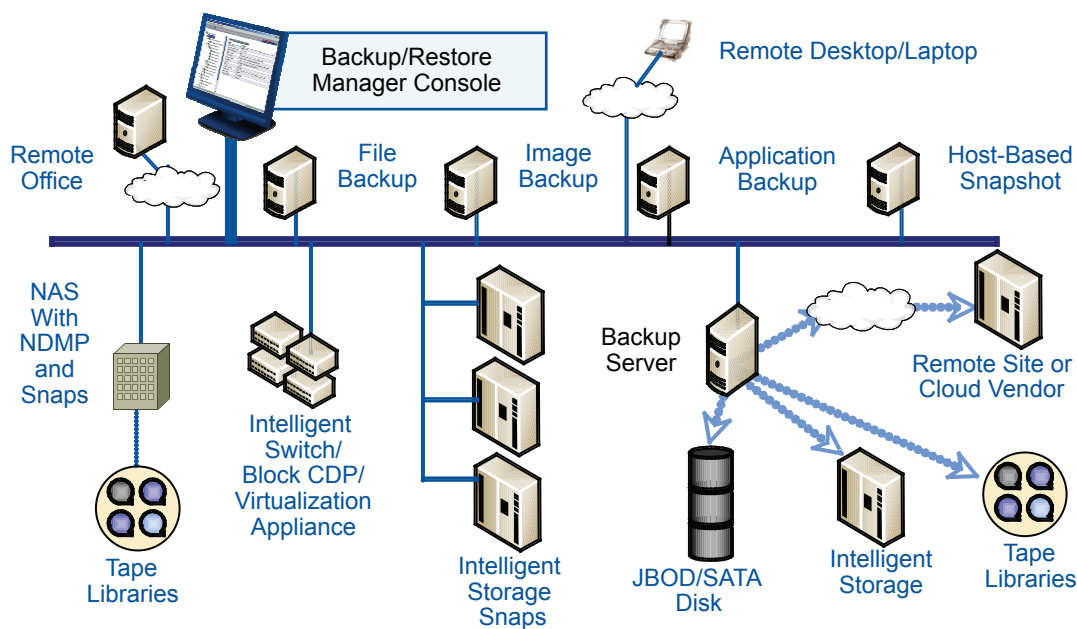
tape and deduplication appliances have gained traction by providing a disk-based solution that requires little change to the backup process. The use of disks for backup and recovery has opened up the market to vendors with products designed specifically to take advantage of disks. Leading backup solutions are capable of protecting data at the file, application, VM and volume levels. These offerings increasingly integrate with a handful of storage arrays and/or NAS filers to catalog snapshots. Some products even offer integration with a replication engine.

Expect traditional backup products to transform into recovery management solutions that may not necessarily own all the data capture and data transfer techniques used. This means that, in

addition to traditional backup and recovery (application, file and image-based, etc.), there will be much stronger support for server-based replication, storage-array-based replication, intelligent-switch and/or network-based replication solutions. Examples include CommVault's Simpana cataloging replication and snapshots from host- and array-based replication solutions, and Symantec NetBackup cataloging NetApp snapshots.

We will have a "manager of managers," a common and established concept in the networking and system management domains, whereby a hierarchy of federated management tools feed into each other, percolating up to an overall unified recovery manager, allowing for more-simplified implementation of several tiers and service levels that offers centralized monitoring, reporting and control (see Figure 2).

FIGURE 2
Unified Recovery Management



Source: Gartner (August 2010)

Action Item: Consider the concept of unified recovery management to be an essential component of your overall backup strategy and architecture, and evaluate incumbent and prospective vendors on their ability to robustly deliver this capability.

3.7 Perform Regular Restore Testing

Backups may be unrecoverable for many reasons. Some of the more common issues are server configuration and application deployment updates, user or operator error in the backup process, and hardware and software failures. In most organizations, backups are initially set up and then automatically run. Backup verification tends

to be only a review of the backup log(s), with a quick scan for failures and error messages. This process may be acceptable for determining whether data was successfully written to the backup media, but it doesn't provide information about whether the data is recoverable, nor does it validate that the right information was backed up. Some businesses have instrumented backup reporting tools to better understand how backups trend over time, and to get more visibility into backup success and failures.

Still, actual recovery is the only way a data center can be certain that data is fully recoverable. Backup/restore testing has

become a dying practice in most data centers, with the end result being that organizations could be far less resilient than they believe they are.

Action Item: Gartner recommends performing data recovery testing at least once a year on a subset of data to ensure that the backup strategy can effectively meet the stated protection SLAs. More-frequent testing of the most mission-critical data may be warranted.

Gartner RAS Core Research Note, G00206028, Dave Russell, 27 August 2010



About Syncsort

Company Overview

Syncsort is a global software company that helps the world's most successful organizations rethink the economics of data. Syncsort provides record-setting performance and rapid time to value through easy to use data integration acceleration and data protection solutions. With over 12,000 deployments, Syncsort has transformed decision making and delivered more profitable results to thousands of customers worldwide. For more information visit: www.syncsort.com.

About NetApp Syncsort Integrated Backup

NetApp Syncsort Integrated Backup (NSB) is a turnkey data protection solution that combines Syncsort data protection software, NetApp Snapshot, clone and replication technology, and high-performance NetApp® disk storage to help deliver fast, reliable, efficient backup and recovery for heterogeneous data centers.

NetApp and Syncsort began their engineering relationship in 1999. Since then, Syncsort continued to certify its software products with NetApp solutions, as well as develop enhanced functionality to align with NetApp's evolving product strategy. In August of 2009, because of their expanding joint customer base, NetApp and Syncsort signed a reciprocal support agreement providing support escalation coverage for joint customers.



In August 2010, NetApp and Syncsort announced a significant expansion of their relationship in the form of NSB, which initiated a new, more proactive sales, marketing and engineering alliance.

Together, both companies are bringing together best-in-class solutions to meet the many challenges of data protection in a continuously connected world.

Source: Syncsort