DATA PROTECTION ON THE ROAD TO ‘ZERO BACKUP’

A variety of relatively new technologies will replace traditional methods of backup and recovery, while at the same time expanding the role of backup data from being an ‘insurance policy’ to being a business asset.

KEY FINDINGS

- Innovative vendors are expanding the role of backup data from a mere insurance policy that protects against failures and outages into a business asset that can be leveraged for additional use cases, such as test and development and data analytics.
- According to end-user studies conducted by 451 Research, 35% of IT organizations plan to redesign their backup and/or disaster-recovery infrastructures over the next year.
- More than half of midsize and large enterprises are using array-based snapshots in conjunction with replication and, in many cases, elements of traditional backup/recovery software, such as cataloging and indexing.
ABOUT 451 RESEARCH

451 Research is a preeminent information technology research and advisory company. With a core focus on technology innovation and market disruption, we provide essential insight for leaders of the digital economy. More than 100 analysts and consultants deliver that insight via syndicated research, advisory services and live events to over 1,000 client organizations in North America, Europe and around the world. Founded in 2000 and headquartered in New York, 451 Research is a division of The 451 Group.

© 2015 451 Research, LLC and/or its Affiliates. All Rights Reserved. Reproduction and distribution of this publication, in whole or in part, in any form without prior written permission is forbidden. The terms of use regarding distribution, both internally and externally, shall be governed by the terms laid out in your Service Agreement with 451 Research and/or its Affiliates. The information contained herein has been obtained from sources believed to be reliable. 451 Research disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although 451 Research may discuss legal issues related to the information technology business, 451 Research does not provide legal advice or services and their research should not be construed or used as such. 451 Research shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The reader assumes sole responsibility for the selection of these materials to achieve its intended results. The opinions expressed herein are subject to change without notice.

New York
20 West 37th Street, 6th Floor
New York, NY 10018
Phone: 212.505.3030
Fax: 212.505.2630

San Francisco
140 Geary Street, 9th Floor
San Francisco, CA 94108
Phone: 415.989.1555
Fax: 415.989.1558

London
Paxton House (5th floor), 30 Artillery Lane
London, E1 7LS, UK
Phone: +44 (0) 207 426 0219
Fax: +44 (0) 207 426 4698

Boston
125 Broad Street, 4th Floor
Boston, MA 02109
Phone: 617.275.8818
Fax: 617.261.0688
TABLE OF CONTENTS

SECTION 1: EXECUTIVE SUMMARY 1
1.1 INTRODUCTION ........................................... 1
1.2 KEY FINDINGS ........................................... 1
1.3 METHODOLOGY ........................................... 2

SECTION 2: WHAT ARE THE PROBLEMS WITH BACKUP AND RECOVERY? 3
2.1 USERS WEIGH IN ........................................... 3
   Figure 1: Top Backup Pains ............................. 3
2.2 DATA-PROTECTION REDESIGN ......................... 4
   Figure 2: Storage Projects – Time Series of Top Categories .... 4
   Figure 3: Time Distribution of Storage Staff ........... 5

SECTION 3: ARRAY-BASED SNAPSHOTS AND REPLICATION 7
3.1 WHAT IS A SNAPSHOT? ................................... 7
   Figure 4: Most Replication Occurs at the Array Level ... 8
3.2 SNAPSHOTS + REPLICATION + BACKUP ................ 8
   Figure 5: Primary Method of Backup. ................ 9
3.3 BENEFITS ............................................... 10
3.4 DRAWBACKS ........................................... 11
3.5 PLAYERS AND RECENT PRODUCT DEVELOPMENTS .... 11
   3.5.1 Disk-Array Vendors. ............................. 14

SECTION 4: COPY DATA VIRTUALIZATION/MANAGEMENT
   AND CONVERGED DATA MANAGEMENT 18
4.1 COPY DATA VIRTUALIZATION ............................ 18
4.2 COPY DATA MANAGEMENT ............................... 19
4.3 CONVERGED DATA MANAGEMENT .................... 20

SECTION 5: RECOMMENDATIONS FOR VENDORS AND END USERS 22
5.1 RECOMMENDATION FOR VENDORS ..................... 22
5.2 RECOMMENDATIONS FOR END USERS ................ 22

INDEX OF COMPANIES 23
SECTION 1
Executive Summary

1.1 INTRODUCTION
End-user studies conducted by 451 Research have shown that IT managers and administrators are fed up with traditional methods of data protection. Problems include the inability to complete backups within allotted ‘backup windows,’ failure to meet increasingly stringent recovery time objectives (RTO) and recovery point objectives (RPO), slow data access and retrieval times, and exorbitant costs, to name just a few. Of course, these are not new issues and vendors have in the past tried a variety of technologies in an effort to diminish some of the problems, but all of those attempts have fallen short of truly eliminating end users’ problems with backup and recovery.

More recently, some vendors have begun to address these problems by rethinking the ‘old school’ approaches to data protection and moving toward techniques that can eliminate backup windows while slashing RPO/RTO and providing near-instant access and retrieval times. The ultimate goal is to achieve a ‘zero backup’ strategy that eliminates the traditional backup and recovery processes while still providing full data protection.

In our view, some of the more promising technologies in this context are array-based snapshots coupled with replication and elements of traditional backup/recovery software (covered in Section 3 of this report), as well as related technologies such as copy data virtualization, copy data management and converged data management (covered in Section 4).

In short, the days of nightly incremental backups and weekly full backups are fading fast. It’s time that primary storage, via its snapshot technology, plays a bigger role in end users’ data-protection strategies.

1.2 KEY FINDINGS
• Innovative vendors are expanding the role of backup data from a mere insurance policy that protects against failures and outages into a business asset that can be leveraged for additional uses cases, such as test and development and data analytics.

• According to end-user studies conducted by 451 Research, 35% of IT organizations plan to redesign their backup and/or disaster-recovery infrastructures over the next year.

• Primary storage systems are taking on an increasingly important role in the data-protection hierarchy.

• 451 Research studies show a steady trend toward IT organizations using array-based snapshots as a foundation for their data-protection strategies. Meanwhile, the use of traditional backup/recovery software alone is steadily decreasing.
• More than half of midsize and large enterprises are using array-based snapshots in conjunction with replication and, in many cases, elements of traditional backup/recovery software, such as cataloging and indexing.

• We expect significant innovation in the data-protection space over the next three to five years, from both startups and the established incumbents, driven by end users’ demands for deriving more business value from backup data.

1.3 METHODOLOGY

This report on data-protection trends in the IT industry is based on a series of in-depth interviews with a variety of stakeholders in the industry, including technology vendors and IT storage professionals at end-user organizations across multiple industries. This research was supplemented by additional primary research, including attendance at trade shows and industry events.

Reports such as this one represent a holistic perspective on key emerging markets in the enterprise IT space. These markets evolve quickly, though, so 451 Research offers additional services that provide critical marketplace updates. These updated reports and perspective are presented on a daily basis via the company’s core intelligence service, 451 Market Insight. Forward-looking M&A analysis and perspectives on strategic acquisitions for technology companies are also updated regularly via 451 Market Insight, which is backed by the industry-leading 451 Research M&A KnowledgeBase.

Emerging technologies and markets are also covered in additional 451 channels, including Datacenter Technologies; Storage; Enterprise Platforms & Infrastructure Software; Networking; Information Security; Data Platforms & Analytics; Development, DevOps & Middleware; Social Business Applications; Service Providers; Cloud & IT Service Markets; European Services; Multi-Tenant Datacenters; Enterprise Mobility; and Mobile Telecom.

Beyond that, 451 Research has a robust set of quantitative insights covered in 451 products such as ChangeWave, TheInfoPro, Market Monitor, the M&A KnowledgeBase and the Datacenter KnowledgeBase.

All of these 451 services, which are accessible via the Web, provide critical and timely analysis specifically focused on the business of enterprise IT innovation.

This report was written by Dave Simpson, Senior Storage Analyst. Any questions about the methodology should be addressed to dave.simpson@451research.com.

For more information about 451 Research, please go to: www.451research.com.
SECTION 2
What Are the Problems With Backup and Recovery?

2.1 USERS WEIGH IN

According to end-user studies conducted by 451 Research, most of the pains associated with current methods of backup and recovery spring from the sheer explosion of capacity in most IT organizations, particularly in ‘big data’ and other huge database environments. Directly related to data growth is the inability to get all of the data backed up within the allotted backup window. Rapid adoption of virtualization technology and the explosion in unstructured content have only exacerbated these problems. And as shown in Figure 1, results from the latest Storage study from TheInfoPro, a service of 451 Research, indicate these problems are escalating.

FIGURE 1: TOP BACKUP PAINS
Source: 451 Research’s TheInfoPro Storage Study – Wave 18

Q. What are your top backup pains?

Exceeding Backup/Recovery Window 46% (67%)
Data Growth 18% (56%)
Managing Backup Hardware/Software 21% (27%)
Tape Management 11% (20%)
Defining a Retention Policy 18% (16%)
Move to Snapshot-based Backup 4% (8%)
Low De-duplication Ratio 7% (5%)
Other 8% (18%)

Preliminary results from the 2015 study, which is currently in progress, show that these problems will become even worse over the next year.

Based on our interviews with IT storage managers and administrators, as a result of the aforementioned problems companies are finding it extremely difficult – often impossible – to meet their increasingly stringent RTO, RPO and SLA requirements. As such, it’s no surprise that storage administrators point to backup hardware/software management as a key pain point (cited by 27% of the respondents in Figure 1).
2.2 DATA-PROTECTION REDESIGN

As a result, IT and storage managers are more willing than ever to redesign their data-protection processes, which in many cases means replacing their existing backup/recovery software vendor – sometimes in favor of relatively unknown startups.

Here are some representative comments from IT storage managers who participated in the Storage study:

“The backup window today exceeds our limits.” – Large enterprise, materials/chemicals industry

“To date, we’ve been able to respond to the high growth rate for backup. [However], we will need additional [data protection] functionality with our virtualization effort.” – Large enterprise, education industry

“We’re reducing the use of [traditional backup/recovery software] and using snapshots to do backup.” – Large enterprise, education industry

In the study of midsize and large enterprises, 22% of participants say backup redesign is a storage project priority, and another 13% say redesigning their disaster-recovery (DR) procedures is a priority, as shown in Figure 2. If we combine those two responses, we see that 35% of the IT organizations plan to redesign their data-protection procedures and infrastructure.

**FIGURE 2: STORAGE PROJECTS – TIME SERIES OF TOP CATEGORIES**

*Source: 451 Research’s TheInfoPro Storage Study – Wave 18*

Q. What are your organization’s top storage-related projects in the next 12 months?

- Technology Refresh and Capacity Expansion
- Backup Redesign
- DR Redesign
- Consolidation
- Flash Implementation
- New/Replatformed Applications
- Cloud Storage
- Moving Datacenter
- Efficiency and Cost Reduction
- Archiving
From the end users’ perspective, the problems associated with existing approaches to backup and recovery translate into much too much time being spent on backup. In fact, in the Storage study, backup administration was cited as the number one time drain out of all storage tasks, as shown in Figure 3.

**FIGURE 3: TIME DISTRIBUTION OF STORAGE STAFF**

*Source: 451 Research’s TheInfoPro Storage Study – Wave 18*

Q. What percentage of time does your staff spend performing each of the following functions?

It's clear that it's time for a change, and IT organizations are more than willing to consider alternative technologies – even if it means a rip-and-replace overhaul of existing backup infrastructure and procedures.

The rapid adoption of virtualization has only exacerbated the data-protection problems, giving rise to backup/recovery startups that specialize in protecting only virtualized environments. But regardless of whether IT organizations are trying to protect virtual or physical servers, the problems with existing backup-and-recovery methods are essentially the same.

Although technologies such as incremental backups, data de-duplication and compression have to a degree addressed some of the problems associated with traditional approaches to backup, they haven’t solved the (more important) issue of recovery. In the case of data de-duplication, for example, data often has to be rehydrated for recovery, which can negatively affect the ability to meet RTO requirements.
What are end users doing to alleviate some of the problems with existing methods of data protection? In midsize and large enterprises, we see increased interest in using disk array snapshots as the underpinning for data-protection strategies, which we examine in the next section of this report. In fact, many IT managers and vendors alike think that array-based snapshots will be the only way to solve data-protection problems in big data environments because traditional methods of backup and recovery simply can’t handle the capacity levels found in these environments.

In addition to increased use of array-based snapshots, another related trend that is rapidly gaining interest among midsize and large enterprises is referred to as copy data management or copy data virtualization, which we cover in Section 4 of this report.
SECTION 3
Array-Based Snapshots and Replication

3.1 WHAT IS A SNAPSHOT?

Most of the newer, more innovative approaches to data protection have their roots in snapshots. Unless otherwise noted, the term ‘snapshots’ in this report refers to disk-array-based snapshots, which are sometimes referred to as hardware-based snapshots, SAN-based snapshots or storage snapshots. This is in contrast to, say, hypervisor-based snapshots, which are used by all backup applications. And some use the term ‘snapshots’ in an even looser fashion, denoting any copy of data.

In fact, snapshots are not true copies of data (even though they’re often referred to as ‘point-in-time copies’). Rather, they are ‘virtual copies’ that keep track of changes made to a base copy of a volume, file or file system over time and then apply those changes in the order they were captured to re-create the data as it was at a specific point in time, typically by using metadata pointers or reference markers. Essentially, snapshots allow users to roll back a file, application, VM, etc., to a previous point in time. As such, they are somewhat analogous to a table of contents. Disk-array vendors vary in how they implement snapshots (e.g., copy-on-write, redirect-on-write, split mirrors, etc., each with performance-capacity trade-offs), but all of the snapshot methods can be used efficiently in the context of data protection.

As we review later in this section, the key benefits of using snapshots, compared to traditional backup software, are significantly faster backups and, more significantly, restores – in part because the snapshots are resident on high-performance primary storage systems. In addition, most snapshots techniques are very space efficient.

There’s nothing new about using snapshots as a form of data protection or, more accurately, point-in-time recovery. However, snapshots alone do not constitute data protection. For one, they reside on the same array as the production data; as such, in the case of a failure on (or loss of) that array, the data is not protected. This is why most IT organizations combine snapshots with off-site replication for full data protection.

Although a number of alternatives exist – including hypervisor-based replication – the majority of replication still takes place at the array level, as shown in Figure 4.
FIGURE 4: MOST REPLICATION OCCURS AT THE ARRAY LEVEL

Source: 451 Research’s TheInfoPro Storage Study – Wave 18

Q. Of the replication occurring in your environment, what percentage is occurring at the following levels?

- Application-level Replication: 31%
- Database-level Replication: 37%
- Host-level Replication: 24%
- Hypervisor-level Replication: 44%
- Array-level Replication: 76%

3.2 SNAPSHOTS + REPLICATION + BACKUP

But just combining array snapshots and replication still doesn’t provide the functionality that users expect from traditional backup applications. What’s new in this space is that ISVs – as well as a few disk-array vendors – are combining snapshots and replication with elements of traditional backup/recovery software, such as indexing, cataloging, scheduling and, if needed, tape-based backup.

All of the leading disk-array vendors, as well as most of the newer players in the all-flash and hybrid-array space, have tools that provide a degree of snapshot management. Although the disk-array vendors continue to improve their snapshot management tools, there are a number of potential drawbacks to using the native snapshot management tools from those vendors. For one, until recently they were proprietary; in other words, snapshot management tools could only be used from Vendor A with disk arrays and snapshots from Vendor A (although there are now some exceptions, as discussed in Section 3 of this report).
Using proprietary snapshot management tools for data protection can be extremely complex and time-consuming in heterogeneous environments that require tools from different vendors. Another drawback is that the array vendors’ snapshot management tools are often not tightly integrated with backup/recovery applications (although, again, this is slowly changing), and the array vendors’ tools typically do not have features such as centralized indexing, cataloging and scheduling. In addition, these tools often require a good deal of scripting – a task so complex that it sometimes requires dedicated personnel, particularly in large heterogeneous environments. And, of course, snapshot management scripts are not portable across disk arrays from different vendors.

For these reasons, many IT organizations have turned to third-party backup/recovery software vendors to augment their array snapshot management and to provide true data protection based on those snapshots.

451 Research has been tracking this trend for several years, and our research shows a steady trend away from using traditional backup/recovery applications without snapshot management to using a combination of snapshots, replication and backup/recovery applications (see Figure 5).

**FIGURE 5: PRIMARY METHOD OF BACKUP**
*Source: 451 Research’s TheInfoPro Storage Study – Wave 18*

Q. What is your primary method of backup?

As shown in the chart, among midsize and large enterprises, the percentage of IT organizations that rely solely on traditional backup software – without using array snapshots – dropped from 62% in 2013 to 44% in 2014. Meanwhile, the percentage of users that rely on snapshots and replication followed by backup software increased from 24% to 31%. And the percentage of organizations that use snapshots and replication managed and/or cataloged by backup software doubled.
from 10% to 20%. Only 5% of the IT organizations in the 2014 study rely solely on snapshots and replication (i.e., without backup software for management of the snapshots) – this is a practice we don’t recommend because snapshots alone do not provide the full data protection that backups provide (as discussed below).

Preliminary data from our 2015 study of IT storage professionals indicates that the trend illustrated in Figure 5 will continue over the next several years.

Although there is a variety of different techniques for taking snapshots (e.g., copy-on-write, redirect-on-write, split mirrors, etc.), snapshots essentially track blocks of data that have been changed and can update them to a copy of the data either continuously or at user-defined intervals (e.g., every 15 minutes, once per hour). This is typically done by using metadata and pointers, which make snapshots very space-efficient. Among disk-array vendors, snapshot techniques vary in terms of performance, efficiency, capacity utilization and the performance effect on primary storage during the snapshot process, but all snapshot methods can be used effectively in the context of data protection.

How often snapshots are taken depends largely on the organization’s RPO, which has to be balanced against capacity utilization (and cost) and, in some cases, primary storage performance degradation.

3.3 BENEFITS

Combining array snapshots with elements of traditional backup/recovery software essentially combines the best of both: Snapshots enable rapid (almost instantaneous) recovery, while backup software provides affordable retention, catalog-based searches and restores.

The primary benefit of adding array snapshots to a data-protection strategy is very fast restore/recovery times, typically measured in seconds. This is in part because the snapshot data is stored in native format (as opposed to a proprietary backup format). Having near-instant restore times – coupled with point-in-time recovery – via snapshots solves two of end users’ biggest pain points with data protection: the inability to meet RTO and RPO. This in turn allows IT managers, backup/recovery administrators and application owners to better meet SLAs. Although users still need to do off-disk backup of the snapshots for full data protection, they can recover from the snapshots directly – without the need to first perform full restores.

Another benefit of using array-based snapshots for data protection is that, unlike traditional backup applications, snapshots have minimal or no impact on the performance of production servers. Yet another potential benefit of using application-aware snapshots is that they enable application owners (e.g., database administrators, VM administrators) to handle backup and recovery tasks, thus reducing backup-specific staffing requirements.

Although this report focuses primarily on the data-protection benefits of using array-based snapshots, it’s important to note that snapshots can also be used effectively for other use cases, such as test/development and analytics, via ‘live clones’ or ‘virtual copies.’
3.4 DRAWBACKS

As shown in Figure 5, more than half of the IT organizations in midsize and large enterprises use hardware-based snapshots to some degree. So why don’t all IT organizations base their data-protection strategies on array-based snapshots? One reason is that it can be expensive; depending on the type of snapshots used by the disk-array vendor, snapshots can consume a lot of primary storage capacity, which translates into high costs. The cost problem can be compounded if users need to retain large amounts of snapshots. In addition, snapshot management tools have sometimes been sold separately from the arrays and/or data-protection applications, which also contributed to high cost.

Another reason why array-based snapshots have not been used by a higher percentage of IT organizations is sheer complexity, particularly in the case of using snapshot management tools from different array vendors. To add to the complexity, using snapshots as the foundation for data-protection strategies often requires coordination between different IT teams – for example, administrators/managers of primary storage, applications, backup/recovery and virtual environments.

However, some of the recent developments in this space are driving down both the complexity and costs that have been historically associated with snapshot-based data protection. Many disk-array vendors are now bundling snapshot technology and snapshot management functionality for free with their arrays, further reducing costs.

3.5 PLAYERS AND RECENT PRODUCT DEVELOPMENTS

This section covers some of large vendors that are leveraging array-based snapshots for data protection, with a focus on recent product developments. The flurry of product announcements on the vendor side over the last year or so is reflective of the heightened end-user interest in using hardware snapshots in backup/recovery scenarios.

Commvault

Along with OEM partner NetApp, Commvault was an early advocate of using array snapshots, coupled with array-based replication and traditional backup/recovery technologies, as a foundation for data-protection strategies. The Commvault-NetApp partnership resulted in NetApp’s SnapProtect software, which is essentially an OEM version of Commvault’s Simpana software and IntelliSnap technology that is optimized specifically for NetApp environments. (We cover NetApp’s SnapProtect below.)

Commvault bundles IntelliSnap technology into the high-end licenses for the Simpana suite, but IntelliSnap is also available as a separately priced stand-alone product for users that are only interested in array snapshot management.

Simpana IntelliSnap technology centralizes snapshot management across heterogeneous arrays, automates application recovery and links snapshots to backup processes. Key features include snapshot catalog, index and search functionality, as well as backup to tape.
Although it would appear that Commvault’s IntelliSnap competes with NetApp’s (IntelliSnap-based) SnapProtect, it’s important to note that SnapProtect can only be used with NetApp arrays and software, whereas IntelliSnap can be used in heterogeneous environments (that may or may not include NetApp arrays and software). Early deployments of IntelliSnap were usually for single-vendor array environments (usually NetApp), but Commvault says that recent deployments are more likely to be in heterogeneous array environments (although in most cases they’re limited to arrays from only two vendors).

We consider Commvault to be a leader in the movement toward using array-based snapshots for data protection for a number of reasons, including its early entry into the market, its OEM relationship with NetApp (as well as Hitachi Data Systems), application-aware features and breadth of compatibility, and support for a wide variety of disk arrays from different vendors. (In contrast, most other ISVs addressing snapshot-based data-protection support one vendor’s arrays, and usually it’s NetApp.)

Commvault introduced support for array-based snapshots in 2010 with the release of its Simpana 9 platform. With the release of Simpana 10 in 2013, Commvault’s IntelliSnap supported most disk arrays from NetApp, EMC, IBM, Hitachi Data Systems, HP, Dell, Fujitsu and Nimble Storage. In 2014, Commvault added support for arrays from Huawei and Oracle to the platform. It also added support for EMC’s VPLEX (although only in configurations with VMAX arrays behind VPLEX), as well as DataCore Software’s virtualization software, which virtualizes heterogeneous arrays.

We expect Commvault to broaden its compatibility matrix in the Version 11 release of Simpana – which is due in the fourth quarter of 2015 – with support for array snapshot technology from vendors in the product categories of all-flash arrays, converged platforms and hyperscale storage systems. Also with the Simpana 11 release, we expect Commvault to add support for other array vendors’ replication technologies; currently, Simpana’s snapshot management tools only support NetApp’s replication technologies.

Commvault supports disk-array partners through its IntelliSnap Connect Program. In some cases, Commvault does the heavy lifting on integration, while in other cases it supplies developer kits to vendors that in turn do most of the integration work.

NetApp

As mentioned previously, Commvault’s work on NetApp’s array-based snapshots led to an OEM partnership in 2011 that in turn led to NetApp’s SnapProtect software. SnapProtect is essentially an OEM version of elements of Commvault’s Simpana that is optimized for NetApp-only environments. Although SnapProtect has been available for almost four years, adoption has just begun to pick up. NetApp attributes the slow adoption in part to the fact that SnapProtect is a full-fledged backup/recovery application that in many cases requires customers to ditch their existing backup application(s), which many IT organizations are reluctant to do. (SnapProtect only works with NetApp’s arrays and software; as such, many NetApp shops use SnapProtect for NetApp systems and third-party backup/recovery software for heterogeneous systems.) However, adoption of SnapProtect is ramping up. For example, NetApp says that it recently had 100% year-over-year quarterly growth in SnapProtect adoption.

Although NetApp and Commvault pioneered the use of array snapshots for data protection, a number of other ISVs that specialize in backup/recovery have also recognized the need to add support for array-based snapshots to their data-protection portfolio. Examples include backup/recovery market leader
Symantec, Asigra and Veeam, as well as array/backup vendors such as EMC, IBM and HP. (Vendors specializing in copy data management/virtualization also leverage array-based snapshots; these are covered in Section 4 of this report.)

**Symantec**

Symantec’s storage operation (to be spun out as Veritas Technologies later this year) has been focused on array-based snapshots for some time. The company added support for NetApp’s snapshots three years ago with the release of NetBackup 7.5 and the Replication Director feature.

Symantec says that adoption of NetBackup’s array-based snapshot management has been slow in NetApp environments, but attributes the slow adoption in large part to the difficulty (for both NetApp customers and Symantec) of migrating from NetApp’s Data ONTAP 7-Mode to its Clustered Data ONTAP operating environment. With the forthcoming 7.7 release of NetBackup (due this summer), Symantec will fully support Clustered Data ONTAP, and the company expects that to spur adoption of NetBackup Replication Director.

However, Symantec has aggressive plans to support other disk-array vendors’ snapshots for data protection. With the release of NetBackup 7.6.1 in early March 2015, Symantec added support for EMC’s VNX arrays. The company says that support for other vendors’ arrays and snapshots will be phased in, and that by the end of 2015 (about the same time that it is expected to split into two companies) Replication Director might support arrays/snapshots from HP (3PAR StoreServ), Hitachi Data Systems, Fujitsu and Pure Storage, as well as EMC's Isilon arrays and snapshots. Support for IBM and Dell arrays are also on Symantec’s roadmap but are not expected in 2015.

Symantec’s Replication Director is available as a NetBackup option at an extra charge or, in some licensing packages, bundled with NetBackup at no extra charge.

**Asigra**

The fact that large backup/recovery vendors such as Commvault and Symantec are putting significant R&D dollars into managing array-based snapshots is a good indication of interest in the technology, but smaller backup/recovery vendors are also getting into the act. For example, Asigra announced support for NetApp’s snapshots in early 2012 and in late 2012 added a Snapshot Manager feature in its Asigra Cloud Backup software that provided snapshot scheduling and cataloging. Snapshot Manager is bundled with Asigra Cloud Backup at no extra charge.

Asigra is currently limited to only NetApp environments, but it plans to expand its compatibility matrix. For example, support for EMC’s VNX – as well as FreeBSD’s Z File System (ZFS) snapshots – is in the works, although Asigra is not discussing time frames for delivery.

**Veeam**

Veeam Software, which specializes in data protection for virtual environments, is another backup/recovery vendor that is leveraging array-based snapshots for backup. Veeam’s software works with disk arrays/snapshots from NetApp and HP (3PAR StoreServ and StoreVirtual arrays) via two technologies that are integrated into its flagship Veeam Backup & Replication software:
Veeam Explorer for Storage Snapshots (for recovery) and Backup from Storage Snapshots (for backup). Veeam has supported HP arrays/snapshots since 2013, and NetApp’s since late 2014.

Veeam does not provide snapshot management per se (e.g., scheduling, indexing, cataloging), but enables backup/recovery from HP and NetApp snapshots. (We expect Veeam to add support for one other vendor’s arrays and snapshots in the forthcoming Version 9 of Veeam Backup & Replication, which is due in late 2015.)

Unlike some other backup/recovery vendors, Veeam combines VMware’s vSphere hypervisor-based snapshots and Changed Block Tracking (CBT) technology with HP or NetApp snapshots. (Veeam supports all versions of NetApp’s operating software, including Clustered Data ONTAP, as well as SnapMirror replication and SnapVault archiving, but only in VMware environments.) Veeam contends that using a combination of array-based snapshots together with VMware’s snapshots and CBT technology eliminates the need for a proxy ESXi host and results in better performance.

### 3.5.1 DISK-ARRAY VENDORS

So far, this report has focused primarily on what Commvault, NetApp and a handful of ISVs are doing with array-based snapshots in data-protection scenarios. But some of the leading disk-array vendors (all of which are also backup/recovery vendors) have also made some interesting announcements in this space that go beyond their proprietary snapshot management point products or ‘element managers.’ Examples include EMC, IBM, HP and Dell (the other leading disk array vendor – Hitachi Data Systems – is covered in the next section of this report).

**EMC**

EMC has a number of tools, or element managers, that can be used to manage array-based snapshots (e.g., Unisphere, AppSync), but in the context of this report we focus on its NetWorker Snapshot Management software and a new product called ProtectPoint, which illustrates the new role of both primary storage and snapshots in data protection.

NetWorker Snapshot Management was originally designed only for managing snapshots on EMC’s block-based disk arrays (e.g., VMAX, VNX, RecoverPoint). However, in 2014 it added support for EMC file-based arrays (VNX, Isilon), as well as NetApp arrays and snapshots. NetWorker Snapshot Management, like Commvault’s Simpana IntelliSnap, combines elements of EMC’s NetWorker backup/recovery software with the benefits of array-based snapshots for not only data protection, but for other use cases as well, such as test and development.

With the introduction of its VMAX3 arrays in mid-2014, EMC rewrote the VMAX snapshot technology. This in part led to the late-2014 introduction of its ProtectPoint software (covered in detail in the 451 Research report *EMC rethinks data protection with ProtectPoint*).

ProtectPoint put EMC on the road to what it refers to as ‘storage integrated data protection.’ In a nutshell, ProtectPoint enables direct backups between VMAX arrays and Data Domain ‘protection storage’ systems. Direct backups potentially eliminate the need for traditional backup/recovery applications, as well as backup/media servers that typically sit between application servers.
and secondary storage appliances. EMC claims using ProtectPoint improves the performance of backups by 20x. And it's important to note that data is stored in its native format, rather than a proprietary backup application format.

The initial implementation of ProtectPoint works only with VMAX arrays and Data Domain systems, but EMC plans to add support for other arrays in the future (most likely VNX and Isilon next).

The initial version of ProtectPoint lacked the backup/recovery functionality that a traditional backup/recovery application would offer. But in May 2015, EMC announced that it will add support for ProtectPoint to its NetWorker backup/recovery software later this year, thus combining the management functionality of its NetWorker data-protection software (e.g., scheduling, cataloging and policy management) with the performance of snapshots – a central theme of this report.

Also in May 2015, EMC added ProtectPoint support (native) for Oracle Recovery Manager (RMAN), SAP BR*Tools and IBM DB2 Advanced Copy Services. The company also added support for Windows, including SQL Server and Exchange, although support for Windows is not yet native.

**IBM**

As mentioned previously, disk-array vendors’ snapshot management tools are typically used only with arrays and snapshots from that vendor. However, some of the array vendors have expanded their snapshot management tools into heterogeneous environments. For example, IBM’s FlashCopy Manager for snapshot-based data protection was originally developed (15 years ago, although under a different name) for IBM arrays, but Big Blue has for about five years natively supported NetApp arrays and snapshots. Beyond that, FlashCopy Manager’s support for heterogeneous environments depends on the operating system.

For Windows environments, FlashCopy Manager supports virtually all vendors’ disk arrays and snapshots because it leverages Microsoft’s Volume Shadow Copy Service (VSS). For example, FlashCopy Manager can be used with disk arrays from EMC, Hitachi Data Systems and HP, as well as NetApp.

To address array-based snapshot management in Unix/Linux environments, IBM partners with Rocket Software, which develops ‘adapters’ for arrays from a variety of vendors. (IBM signed the partnership deal with Rocket about two years ago.) In the context of IBM’s FlashCopy Manager, Rocket’s adapters support arrays from EMC, Hitachi Data Systems and HP. In VMware environments, IBM’s FlashCopy Manager also supports EMC, Hitachi and HP arrays via Rocket adapters and VMware’s vStorage APIs.

Of course, it’s rare that IBM’s FlashCopy Manager (or any other array vendor’s snapshot management tools) is used with other vendors’ arrays. For example, IBM estimates that less than 10% of its FlashCopy Manager customers are using the software to manage non-IBM arrays in conjunction with IBM arrays. However, as the use of array-based snapshots for data protection becomes more commonplace, IBM expects increased use of FlashCopy Manager in heterogeneous environments (which is why it entered into the OEM partnership with Rocket Software).
FlashCopy Manager can be used in stand-alone mode or in conjunction with IBM's Tivoli Storage Manager (TSM) data protection software. In April 2015, with the release of TSM 7.1.2, IBM added FlashCopy Manager to the TSM Suite for Unified Recovery at no extra charge. (FlashCopy Manager was previously sold separately or bundled with IBM's disk arrays.) Used in conjunction with TSM, which IBM recommends, FlashCopy Manager provides full data protection by offloading backups of snapshots to centralized TSM backup servers and, if required, to tape for long-term retention.

In addition to supporting a number of operating environments and third-party disk arrays, IBM's FlashCopy Manager is application-aware: It supports databases from IBM (DB2), Microsoft (SQL Server and Exchange), Oracle and SAP (including HANA). As of the TSM 7.1.2 release in April, FlashCopy Manager also supports software snapshots on IBM's GPFS clustered file system, which was recently re-branded as Spectrum Scale. Other snapshot-related additions in TSM 7.1.2 include support for HP's 3PAR StoreServ arrays via Rocket Software adapters, and support for IBM's DB2 PureScale snapshots on the Spectrum Scale file system.

**HP**

HP is also putting a lot of R&D into using hardware snapshots for data protection. In January 2015, the company shipped a product called StoreOnce Recovery Manager Central (RMC), which leverages array snapshot-based backup via HP's Express Protect feature. StoreOnce RMC enables direct backups (aka 'flat backups') of HP 3PAR StoreServ arrays to HP StoreOnce backup target systems. This eliminates the need for backups to go through dedicated backup/media servers and traditional backup applications. The company claims backup performance increases by 17x. StoreOnce RMC is part of HP's converged data protection initiative.

StoreOnce RMC uses the HP 3PAR StoreServ SnapDiff technology, which sends only changed blocks to the StoreOnce backup system. Snapshots backed up on StoreOnce are self-contained volumes that can be restored back to any HP 3PAR array in the event of a disaster. The company claims that StoreOnce RMC eliminates the need for scripting.

Key advantages of StoreOnce RMC include fast recovery and elimination of the backup window. In addition, RMC enables application owners and/or VM administrators – rather than backup specialists – to manage snapshots, backup and recovery tasks directly from the VMware vCenter console. HP’s snapshot-based data-protection scheme also leverages the company’s Catalyst software for backup acceleration.

StoreOnce RMC is independent of HP’s Data Protector backup/recovery software, although the company is expected to announce integration of the two products later this year. Also, third-party backup applications could potentially access RMC via open APIs that HP provides. As such, although RMC can eliminate the need for traditional backup/recovery applications, it can also be used in conjunction with those applications.
Today, StoreOnce RMC works only with HP's 3PAR StoreServ and StoreOnce disk arrays – and only in VMware environments. The company says that it plans to support other HP arrays – starting with StoreVirtual arrays later this year – as well as other vendors’ arrays in the future. HP also plans to add support for other hypervisors and specific applications, including Hyper-V and Oracle databases.

**Dell**

Although the company has not announced a product yet, Dell is in the process of integrating its NetVault backup/recovery software with its Compellent disk arrays and snapshots. The integration will be part of the 11.0 version of NetVault, which is due in the second half of 2015.

The NetVault-Compellent combo will initially work with Windows and specific applications such as SQL Server and Exchange, although a future implementation will add support for VMware vSphere, Microsoft Hyper-V and Oracle databases. In addition, Dell says that it will add support for snapshot management on other vendors’ arrays, starting with EMC’s VNX and/or NetApp arrays.

Users will be able to create local snapshot-based backups as well as secondary copies on any backup target supported by NetVault.
SECTION 4

Copy Data Virtualization/Management and Converged Data Management

So far in this report we’ve focused primarily on what disk-array vendors and third-party backup/recovery ISVs are doing with array-based snapshots for data protection. Closely allied with this discussion are the emerging trends of copy data virtualization, copy data management and an even newer technology sometimes referred to as converged data management. Although vendors in this space use different terminologies and technologies, their use cases and value propositions are very similar.

These technologies are closely related to our discussion of snapshot-based data protection because they all utilize snapshots in some way for data protection (and other use cases), but they also address the issue of rapidly proliferating copies of production data. All vendors in these technology categories address data protection to varying degrees, although some are more focused on managing copies of production data and making use of virtual copies for use cases that go beyond data protection, such as application test/development and data analytics.

4.1 COPY DATA VIRTUALIZATION

Actifio

Six-year-old Actifio coined the term ‘copy data management.’ Because a few other vendors have leveraged the term since then, Actifio now prefers to use the term ‘copy data virtualization.’ This change came about in part to distance itself from those vendors and in part to emphasize its key differentiator (the ability to virtualize data).

We’ve covered Actifio’s technology in a number of vendor-specific reports on the company, but in the context of this report, here’s a quick overview of how Actifio’s technology works.

The company’s software captures changed data at the application layer using familiar tools such as Microsoft’s VSS and VMware’s snapshots. However, unlike most backup/recovery applications, Actifio captures data in its native application format, as opposed to a proprietary backup format. Changed data can be captured either out-of-band over an IP network or in-band via Fibre Channel.

Data is managed by Actifio’s appliances (physical or virtual) via the company’s Virtual Data Pipeline (VDP) technology, and is stored on Actifio’s storage systems. (Alternatively, customers can store data on third-party disk arrays.) VDP maintains and manages two data pools: a snapshot pool and what Actifio calls a DeDup pool. The snapshot pool is used to provide instantaneous access to data (at any point in time) via virtual copies, and the DeDup pool uses data de-duplication and compression for capacity optimization. VDP manages the data copies between the two pools to optimize access performance while minimizing capacity utilization.
Actifio does not use array-based snapshots on the primary production disk arrays; rather, it creates its own copy-on-write snapshots. For off-site data protection, Actifio has a number of replication options.

In the company’s early days, customers bought the technology primarily for data-protection purposes. In these instances, Actifio replaced existing backup/recovery software, backup appliances, de-duplication targets, replication software and other data protection tools. Although data protection is still the primary use case for Actifio’s copy data virtualization technology, other use cases subsequently emerged – most notably test/development and data analytics.

### 4.2 COPY DATA MANAGEMENT

**Hitachi Data Systems**

Hitachi Data Systems (HDS) is a relative newcomer to the data-protection market, but its new product – Hitachi Data Instance Director (HDID) – illustrates two of the key trends covered in this report: copy data management and the use of array-based snapshots as the foundation for data protection.

HDID has its roots in HDS’ acquisition of Cofio Software in 2012. We covered the resulting product – then called Hitachi Data Instance Manager – in a [mid-2013 report](#) when HDS introduced the product. The company was quiet for the ensuing two years, but in May 2015 it rolled out the renamed HDID.

HDID has similarities with Actifio’s technology (although for now HDID works only with Hitachi arrays and software), but as of the May 2015 release of version 5 of the product, it also incorporates management of Hitachi’s array-based snapshots in a manner similar to snapshot management tools from array vendors such as EMC, IBM and HP. HDID also integrates with Hitachi’s clone and replication technologies, and works with Hitachi’s block- and file-based arrays.

In addition to snapshot management, features of HDID include ‘live’ backup, continuous data protection, global data de-duplication, replication and archiving (to either the Hitachi Content Platform object store or to Microsoft’s Azure cloud). HDID’s snapshot management encompasses HDS’ Thin Image snapshots, ShadowImage clones and Hitachi NAS (HNAS) director clones. It also integrates with Hitachi’s block-based replication technologies (TrueCopy and Universal Replicator) and HNAS Replication products. HDID works with Windows, Linux and, as of the recent release, AIX platforms.

Although Hitachi says that it is not yet out to replace existing enterprise-class backup/recovery applications, but it will eventually look to replace traditional backup with HDID in what it refers to as ‘storage-based protection.’
Interestingly, the introduction of HDID 5 puts HDS in competition with its data protection partners: Hitachi OEMs Commvault’s Simpana – including IntelliSnap technology – and rebrands it as the Hitachi Data Protection Suite (HDPS). Hitachi also resells Symantec’s NetBackup data-protection software. Hitachi officials claim that, at least for now, the introduction of HDID 5 will not alter its existing relationships with Commvault and Simpana. For more information on Hitachi’s HDID 5, see the 451 Research report *Hitachi Data Systems puts both feet in the data-protection pool*.

**Catalogic**

Catalogic Software describes its ECX platform as ‘copy data management,’ but it does not position ECX as a backup/recovery or data-protection product per se. (Catalogic has a separate, older product called DPX that is a traditional backup/recovery application.)

ECX is a tool for cataloging and managing snapshots. In its first release, it worked only with VMware snapshots and NetApp snapshots and disk arrays. However, in May 2015, Catalogic added support for Microsoft’s Hyper-V and IBM’s Storwize disk arrays.

The key use cases for Catalogic’s ECX are data recovery (but not backup per se), automated disaster recovery (leveraging SnapMirror in the case of NetApp environments), test and development, and analytics. In addition to cataloging and indexing, specific features include search, orchestration, reporting and ‘instant recovery’ via snapshots.

**Delphix**

Delphix is another vendor that uses the term copy data management to describe its products. But like Catalogic, Delphix does not position its platform for data protection (although it does have some technology elements that are similar to Actifio’s, including data virtualization and rapid provisioning). Sometimes referred to as ‘database virtualization,’ Delphix’s technology creates virtual clones of production databases for use cases such as application test and development, reporting, training, and others where data has to be kept in sync with a master database. For more information on Delphix, see the 451 Research report *Delphix’s database virtualization engine now applied to enterprise IT modernization*.

**4.3 CONVERGED DATA MANAGEMENT**

**Rubrik**

One of the more recent entries in the new wave of startups tackling data protection from innovative angles is Rubrik (profiled in the 451 Research report *Rubrik hopes to disrupt the data-protection market: ‘no more backup’*), which launched an early release version of its converged data management platform in April 2015. The product is expected to be generally available in the third quarter of 2015.

Aside from its technology (see below), Rubrik’s lineup of investors and founding executives has already drawn considerable attention to the startup. In addition to VC firm Lightspeed Venture Partners, Rubrik’s investors include John Thompson (Microsoft chairman, CEO of Virtual Instruments and former CEO of Symantec), Frank Slootman (president and CEO at ServiceNow and former CEO at Data
Rubrik received Series A funding of $10m in April.

Rubrik’s founding executives come from companies such as Google, Facebook, Data Domain, Oracle and VMware, and have considerable experience in scale-out file systems and Web-scale architectures.

The company is not divulging details on its technology yet, but some information has been released.

Rubrik sells a combination of commodity hardware appliances and software. In the initial implementation, the appliances are based on 2U Supermicro servers, although Rubrik will qualify other platforms in the future. It claims that users can scale to ‘thousands’ of 10TB or 15TB appliances (which Rubrik calls ‘Briks’) in a Web-scale (petabytes) architecture. Unlike most other data-protection appliances, the Briks combine flash SSDs and rotating HDDs. Another differentiator is that the technology eliminates the need for separate backup/media servers or proxy servers, which other vendors in large backup environments sometimes require.

Like most other backup vendors, Rubrik captures data using VMware’s snapshots and data-protection APIs. The secret sauce is in the management of backup data – including compression, de-duplication and version management – in a Web-scale clustered configuration. The technology stores a single copy of data, but perhaps most importantly, it provides instantaneous (measured in seconds, according to Rubrik) point-in-time recovery of applications/data/VMs for multiple use cases, including data recovery and application test/development. The technology stores data in a single pool, rather than having separate pools for snapshot data and de-duplicated data.

Rubrik claims that its technology can eliminate the need for traditional backup software. For example, an early adopter we spoke to – Langs Building Supplies in Australia – retired its traditional backup software after deploying Rubrik’s technology. (Langs uses Rubrik’s technology in conjunction with hyperconverged systems from Nutanix, and it’s interesting to note that Rubrik cofounder and CEO Bipul Sinha was a founding investor in Nutanix.)

Rubrik is currently limited to VMware virtual environments and Oracle physical environments, but the company plans to expand support to other hypervisors and physical systems in the future. Likewise, the technology currently includes integration with Amazon’s AWS, but Rubrik plans to support other public clouds in the future.

Rubrik’s technology is indicative of another major trend in the data-protection market: deployment simplicity and ease of use. The company claims that its product can be configured and deployed in less than 15 minutes – without professional services engagements, which are typically required by enterprise-class backup/recovery vendors. In addition, Rubrik claims that its technology does not require storage specialists, and that professionals such as VMware administrators or database administrators can perform all data-protection functions.

Although the data-protection space is extremely crowded, making it difficult for startups to enter the market, we expect a few more startups to emerge over the next year, all of which are likely to leverage the technologies and trends covered in this report.
SECTION 5
Recommendations for Vendors and End Users

5.1 RECOMMENDATION FOR VENDORS

• For disk-array vendors, we recommend expanding the support matrix for snapshot management beyond their own arrays. We think that end users are looking to minimize the number of snapshot management applications (ideally, one application) in their data-protection strategies. If the hardware vendors do not expand support to other vendors’ arrays, they will cede this market opportunity to independent ISVs that embrace multiple vendors’ arrays and snapshots.

• Likewise, for third-party backup/recovery software vendors we recommend expanding support matrices to at least the leading vendors’ disk arrays. Commvault has already done this, but other vendors are still limited in their heterogeneous array support.

• Data-protection vendors targeting the midrange and large-enterprise markets should be focused on leveraging array snapshots and managing copy data in use cases that go beyond the backup ‘insurance policy.’ Examples include test/development and data analytics.

5.2 RECOMMENDATIONS FOR END USERS

• Any IT organization that is (a) having trouble backing up data within the allotted backup window, (b) having trouble meeting recovery time or recovery point objectives, or (c) not meeting its data-protection SLAs should be evaluating a data-protection strategy based on the management of hardware snapshots coupled with elements of traditional backup software, such as cataloging, indexing, search and reporting.

• Although using different tools from different vendors to manage snapshots is possible, this approach will eventually become too complex and costly. As such, storage managers and administrators should look for snapshot management tools that can be used in heterogeneous environments in order to minimize the number of management tools, interfaces and vendors they have to deal with.

• Since virtually all midsize and large enterprises suffer from problems related to the proliferation of copies of production data, we recommend evaluating new technologies that are focused on copy data management. Although this may require betting on startups – and overhauling existing infrastructure and procedures – users should bear in mind that the newer vendors have not been saddled with having to adapt existing legacy products to address the new realities of data protection. Just as startup flash vendors have begun to alter the vendor hierarchy in the primary storage arena, it may be time for a new wave of startups to enter the data-protection hierarchy.
INDEX OF COMPANIES

Actifio 18, 19, 20
Asigra 13
Catalogic 20
Commvault 11, 12, 13, 14, 20, 22
DataCore Software 12
Dell 12, 13, 14, 17
Delphix 20
EMC 12, 13, 14, 15, 17, 19
FreeBSD 13
Fujitsu 12, 13
Hitachi Data Systems 12, 13, 14, 15, 19, 20
HP 12, 13, 14, 15, 16, 17, 19
Huawei 12
IBM 12, 13, 14, 15, 16, 19, 20
Microsoft 15, 16, 17, 18, 19, 20
NetApp 11, 12, 13, 14, 15, 17, 20
Nimble Storage 12
Oracle 12, 15, 16, 17, 21
Pure Storage 13
Rocket Software 15, 16
Rubrik 20, 21
SAP 15, 16
Symantec 13, 20
Veeam 13, 14
VMware 14, 15, 16, 17, 18, 20, 21