

An Integrated Approach to Data Reduction: Simplifying Management and Reducing Costs

February 2010

Adapted from Worldwide Data Protection and Recovery Software 2008 Vendor Shares: Market Shows Growth Despite Economic Slowdown by Laura DuBois, IDC #219817 and Worldwide Storage Software 2009–2013 Forecast: Quest for Efficiency Driving Investments in Storage Software by Noemi Greyzdorf, Laura DuBois, and James Baker, IDC #217529

Sponsored by CommVault

Despite the economic downturn, the data protection and recovery software market has fared better than many software markets; IDC predicts that 2009 revenue of nearly \$4 billion (up from \$3.8 billion in 2008) will increase to \$4.9 billion by 2013, representing a CAGR of 4.8%. Growth in the data protection and recovery software market is being driven by investment in online backup services, new data protection options, data deduplication, endpoint and remote branch data protection strategies, the ongoing proliferation of server virtualization, new collaborative applications, and consumer/SOHO segments.

Driving the need for data protection among organizations is the exponential growth in data volume. On average, firms are facing 50–60% data growth annually while dealing with meeting backup windows. Data growth, coupled with the inability to meet backup windows, can drive investments in tape infrastructure, virtual tape libraries, additional backup software licenses, data deduplication, and so forth. In particular, data deduplication to reduce backup capacity, aid backup windows, and reduce network traffic will continue to fuel growth of the data protection market.

To contend with data growth as well as the current economic realities, organizations cannot simply invest in more data protection and recovery software and infrastructure. Increasingly, organizations need to manage their data storage infrastructures more efficiently by intelligently reducing the amount of data they store as well as moving data to the most cost-effective and appropriate tier of storage.

This Technology Spotlight focuses on customer benefits of storage management and data reduction technologies and highlights the role that CommVault plays in this important market.

Improving Storage Management

With data growth soaring and budgets still largely constrained at many organizations, data reduction technology and practices yield significant storage cost and administrative savings. By reducing overall data volume and migrating retained data to the most appropriate tier of storage, organizations can achieve significant reductions in storage costs from several perspectives, including:

- Capacity and thus cost
- Physical footprint
- Power and cooling
- Backup capacity
- Smaller backup windows
- Administrative and management overhead

Efforts to contain storage costs are driving organizations to seek more efficient ways to deploy and manage storage resources. This includes projects focused on consolidation, capacity optimization, and storage tiering. More data and tighter budgets will force IT managers and storage administrators to increase the asset utilization of existing resources and add assets only where they are absolutely needed.

Organizations are implementing numerous storage technologies and practices to improve their storage management efforts. For example, email and file archiving applications provide an efficient approach for managing, storing, indexing, and retrieving email messages and file attachments. Having this capability is critical, especially with the ever-increasing amount of email created by most organizations.

In addition, data deduplication is an important and key storage management technology. Deduplication is an extremely relevant and timely technology that organizations are deploying to reduce data, especially for backup and archive data where there is more redundancy compared with production data on primary storage tiers. As firms evolve their data protection and retention strategies, many are also looking to augment or replace tape with a disk-based approach. The economic advantages of data deduplication are accelerating the evolution of tape from data protection toward long-term data retention.

No meaningful slowdown in data creation, a "save everything" corporate society, and the current economic climate and constrained IT budgets are factors driving IT executives to look for ways to reduce storage capacity requirements and cost. Deduplication as a technology meets this objective.

For many organizations, deduplication is only one part of their storage strategy. By integrating technologies for storage resource management (SRM) — a process for gaining visibility into data scattered across storage systems to optimize storage resources and align data on the most appropriate storage tier — organizations can reduce data, compress backup windows, and also migrate the data that they do retain to the most appropriate and cost-effective storage tier.

The protracted economic crisis has accelerated this shift as companies seek to limit spending on installed environments while investing more in solutions that help them improve asset utilization, boost operational efficiency, and more effectively meet regulatory risk management requirements.

The Data Protection and Software Recovery Market: Definitions

IDC divides the data protection and software recovery market into several different functions that allow organizations to devise various storage management strategies. Email and file archiving applications provide an automated and efficient way of storing, indexing, and retrieving individual email messages and file attachments in real time by individual users, IT staff, and other authorized parties from both inside and outside the firm. Email archiving differs from backup and retrieval solutions. Retrieval solutions are designed to provide regularly scheduled copies of an email server or disk to enable email to be brought back up and running after the server or disk fails. They therefore do not maintain copies of emails exchanged between backups, retain copies of emails deleted by users after the backup is replaced with a newer one, or move email content off primary email servers to more efficient storage systems.

Deduplication technology, which replaces duplicate data with references to a shared unique copy in order to save storage space, is becoming mainstream in today's data protection and archiving environments. More specifically, deduplication refers to any process or algorithm that searches for duplicate data objects (e.g., blocks, chunks, files) and discards those duplicates. When duplicate data is detected, it is not retained; instead, a "data pointer" is modified so that the storage system references an exact copy of the unique data object already stored on disk.

Data protection and recovery software is focused on protection, restoration, and recovery of data in the event of physical or logical errors. Included within the data protection and recovery software market are data protection and continuous data protection (CDP), among other technologies.

Data protection software will schedule a copy of (or continuously copy, if using CDP) a defined data set to tape, disk, or optical devices and is used to recover part or all of the data set if needed because of logical or physical error or site disaster. Included are library and tape media management tools. Although tape is the most common backup medium, many products support backup to disk, and some take full advantage of the increased recovery performance associated with direct-access disk storage. Backup is often used in conjunction with snapshot and data replication software to improve data protection performance.

Deduplication is a feature or component of data protection software. The value of a deduplication feature is reduced capacity, backup windows, and network overhead. In effect, deduplication, when integrated with data protection software, represents a more efficient way to do backups.

Benefits of Storage Management

IDC research continues to show that customers are generating record amounts of data every year, leading them to purchase increasing amounts of storage capacity, even as storage hardware prices continue to decline. Faced with this capacity growth, customers are trying to manage their storage infrastructure more efficiently. This is leading to greater adoption of storage management and data deduplication technology as well as email and file archiving solutions. These new data protection approaches are optimized and well suited for virtual machine environments. These organizations understand that they need to manage the corporate information stored on these devices more effectively, as a factor in their competitiveness in the marketplace and to operate within given budgets and IT resources.

Customers employing a sound email and file archiving strategy benefit from more efficient backup. As a result, more data is moved to secondary storage repositories, easing the amount of data that would be put on costly, primary storage systems.

Deduplication is a key technology innovation that is becoming ubiquitous in the storage industry at large. End-user production sites are gaining significant storage spending, space, backup window, and power/cooling savings from utilizing the technology for their backup and archive data.

Many organizations that have implemented email and file archiving as well as data reduction technology have experienced significant business impact, including the following:

- Reduced storage capacity needs
- Operational cost savings
- Reduced management and administrative costs
- Ability to leverage existing resources
- Ability to protect investments in existing technologies

Many organizations have opposing requirements; they have to back up increasingly more data in limited backup windows while also reducing storage infrastructure and cost. Implementing and integrating technology that can manage storage needs, handle archiving, and dedupe data will enable organizations to achieve measurable benefits. However, organizations must not only reduce the amount of data managed but also identify data that can be moved from more expensive primary storage to lower-cost secondary storage.

Reducing costs without compromising backup and data recovery requirements is a benefit of paramount importance in today's economic climate.

Market Trends

The data protection and recovery software market remains the largest segment in the storage software market. Opportunity in the delivery of cloud-based services for backup and recovery continues to be a growth driver in the overall data protection and recovery software market. Major players and emerging suppliers alike continue to roll out offerings. Areas of investment in the mature data protection and recovery software market include disk-based backup, data deduplication, cloud services, server and desktop virtualization, endpoint and remote office protection, and consumer and SOHO segments.

The drivers in data protection and recovery investments include online backup services; compressed service-level agreements; deduplication adoption; server virtualization proliferation; data protection at the edge; control of the endpoint; desktop virtualization; lower-cost, high-capacity drives for protection and replication; data growth and backup windows; and investments by SMBs.

Product Profile

CommVault is a provider of data management software and services based in Oceanport, New Jersey. The company aims to help customers lower costs and reduce the risks associated with data management by using a single unified architecture for its data protection, archive, replication, resource management, and search software products. Rather than requiring a firm to purchase and deploy a series of point products and acquired technologies, CommVault offers customers a single architectural framework (Simpana) upon which they can deploy one or many different data management applications. These applications share a set of underlying service components for tasks such as policy enforcement, data movement, classification, encryption, reporting, and cataloging. The benefits of this architecture include lower TCO, improved management, and greater deployment flexibility. Simpana 8 customers have been able to reduce their tape archives by up to 90% and reduce data management costs by up to 40%. CommVault's singular architecture allows firms to focus on strategic imperatives, including the trends previously mentioned.

Over the past few years, CommVault has grown from an emerging company into a growth-oriented publicly traded data and information management software pure-play with a yearly revenue run rate of \$280 million and market capitalization in excess of \$900 million. Simpana 8, which was released in January 2009, has expanded deduplication features with block-level deduplication and the ability to store dedupe data on tape. CommVault continues to differentiate itself on its singular approach to data management, where different modules for functions such as backup, archiving, replication, storage resource management, deduplication, and search can be deployed discretely but are built upon and leverage a single unified code base and common architectural platform.

CommVault Simpana software integrates and embeds deduplication throughout an organization's data infrastructure. The software employs a content-aware deduplication process to more accurately find and reduce common patterns in the data stream across disparate applications, file systems, and data.

CommVault's unified architecture is the underpinning for the following capabilities:

- Data protection and backup and recovery
- Data migration and archiving supporting data reduction and space management, compliance, and electronic discovery
- Replication of data

- Software-based block-level data deduplication
- Creation and management of copies of data
- Storage resource discovery and management
- Enterprisewide search

The above capabilities leverage CommVault's common architecture framework, which the company refers to as its Common Platform. The Common Platform leverages a single underlying code base for CommVault's application modules and options. The CommVault Common Platform includes an indexing engine, a cataloging engine and global database, a policy engine, a data movement engine, and a media management engine. This single-platform architectural design differentiates CommVault's products from those of its competitors, some of which offer similar applications built upon disparate underlying software architectures.

CommVault's Simpana software supports block-level deduplication for backup data sent to disk. Deduped data can then be stored on tape in deduped form using CommVault's SILO storage option. Simpana integrates and embeds deduplication throughout a firm's data management infrastructure. The software employs content-aware deduplication to more accurately find and reduce common patterns in the datastream across disparate applications, file systems, and data. In fact, Simpana 8 helped Texas Tech University reduce backup storage requirements by more than 80%, lower storage costs by an anticipated 30% to 40%, and achieve nearly 6:1 dedupe ratios.

Challenges

CommVault's unified architecture provides customers with simple, integrated, and flexible tools to manage and safeguard their data. However, CommVault must contend with other software and storage vendors offering similar product sets. IDC believes CommVault must educate potential customers about the benefits of its unified or integrated approach to data protection, deduplication, and storage management. Additionally, CommVault is competing with well-entrenched Fortune 1000 IT vendors with significant customer investments. Displacing a competitor or an existing solution is challenging, but in the current economic environment, customers appear to be more willing to look beyond their traditional vendors in search of solutions that could provide better value and prevent a vendor from locking in its solutions. Furthermore, CommVault must overcome the perception among potential customers that it is a purveyor of only backup and data protection software. IDC believes it is sound and prudent advice for CommVault to refine its message to emphasize investment protection, ease of integration, and the measurable benefits that integrated, software-based data deduplication brings to customers.

Conclusion

In today's already constrained datacenters and growing distributed remote and branch office locations, field-proven storage management, archiving, and deduplication technologies are yielding significant storage cost, footprint, and power/cooling savings. As firms look to make use of disk in the data protection path, the use of storage resource management technology, email and file archiving solutions, and deduplication is becoming a default requirement. Data deduplication brings measurable customer benefits by allowing customers to forgo acquisition costs for new storage hardware to accommodate the ever-increasing growth of data and improve backup and recovery objectives.

IDC believes CommVault, with its integrated data management and deduplication capabilities in Simpana 8, offers a good deal of value by coupling data reduction with an integrated management interface, providing customers with the necessary tools to cost-effectively manage their growth and optimize their storage resources.

A B O U T T H I S P U B L I C A T I O N

This publication was produced by IDC Go-to-Market Services. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Go-to-Market Services makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

C O P Y R I G H T A N D R E S T R I C T I O N S

Any IDC information or reference to IDC that is to be used in advertising, press releases, or promotional materials requires prior written approval from IDC. For permission requests, contact the GMS information line at 508-988-7610 or gms@idc.com. Translation and/or localization of this document requires an additional license from IDC.

For more information on IDC, visit www.idc.com. For more information on IDC GMS, visit www.idc.com/gms.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com