

..... **A CommVault® Business-Value White Paper**

Unlocking the Value of Global Deduplication for Enterprise Data Management

*Reduce Costs, Reduce Complexity,
Reduce Recovery Times*

Contents

Executive Summary.....	3
Redundant Data Reduces Storage Efficiency.....	3
Challenges in Deploying a Deduplication Solution.....	3
Candidates for Deduplication.....	4
Features To Look For in a Deduplication Solution.....	5
CommVault® Simpana® Software— A Complete Data Management and Deduplication Solution.....	5
Simpana Software Benefits.....	7
Conclusion.....	8
About CommVault.....	8



Executive Summary

If you are like most IT executives, at the same time you are grappling with how best to manage your rapid data growth, you are also faced with the need to keep expenses down. Yet, the thought of improving your data management and protection processes conjures up images of increasing backup storage, networking devices and data center operations—all expensive propositions.

Fortunately, deduplication technology can deliver the data management benefits you require, without negatively impacting your bottom line. Deduplication can revolutionize the way you store and protect your data. But beware. Not all deduplication products deliver a complete solution. Many approaches result in performance bottlenecks, hardware vendor lock-in, incomplete backup workflows, and insufficient data reduction.

The best approaches, such as CommVault® Simpana® software, include deduplication as an integrated component of a complete data management solution.

The business benefits from deduplication start with lower data protection costs and lead to even better recovery times. Data deduplication lets you significantly reduce the amount of disk you need for backup across all storage tiers. With reduced acquisition costs—and reduced power, space, and cooling requirements—your disk becomes suitable for primary backup, restore and retention that can easily encompass months of data. With data on disk, restore service levels are higher, media handling errors are reduced, and more recovery points are available on fast recovery media. The cumulative result is that deduplication-based cost reductions and storage efficiency free up your headcount and resources for more strategic tasks and dramatically speed recovery from any storage tier.

Redundant Data Reduces Storage Efficiency

Data redundancy is a primary contributor to the explosive growth in data. Initially, deduplication focused on eliminating data redundancy in specific cases like full backups, e-mail attachments, and VMware OS images. Over time, however, customers have noticed the pervasiveness of duplicated data. Test and development data multiplies across an organization; replication, backup, and archiving create multiple data copies scattered across your enterprise; and sometimes users simply copy data to multiple locations for their own convenience. Studies estimate that multiple copies of data now require organizations to buy, use, and administer two to 50 times more storage than they would actually need with deduplication. Given that impact on the bottom line, organizations are recognizing that, far from being a niche technology, deduplication needs to become an integrated and mandatory element in their overall IT strategy.

Challenges in Deploying a Deduplication Solution

Like disk-to-disk backup or server virtualization, deduplication should not be evaluated as an isolated product or feature; customers must consider the broader implications of deduplication within the context of their entire data management and storage strategy. Common challenges in deploying a deduplication solution are related to performance, increased complexity of management, and proliferation of deduplicated data silos.

Performance

- Finding and eliminating redundant data can be extremely expensive for an appliance-based deduplication solution. Without contextual knowledge of the data that it deduplicates, it faces significant challenges scaling to the size of most enterprises.

- Storage systems perform best when data is written sequentially, thus smaller data sets can cripple disk performance for most deduplication solutions. By “rote” sharing small segments of data across multiple objects, appliance-based deduplication can lead to widespread data fragmentation. Over time, read, write, backup, and replication performance on deduplication appliances becomes painfully slow.

Increased Complexity of Management

- Many deduplication solutions today behave as if the entire workflow revolves around their value; it is almost impossible to move data from a deduplication appliance to tape in a D2D2T workflow. To reap the benefits of network optimization, an organization needs to either install new hardware or software in its remote offices.
- Many deduplication solutions require organizations to integrate a unique combination of hardware and software, or buy new standalone appliances that must be managed manually on an individual basis. Extra management complexity degrades your storage and network savings, especially as you scale the scope of the amount of data being deduplicated.

Islands of Deduplication

- Proprietary solutions create vendor lock-in, combining poor performance with proprietary storage layouts. They make it nearly impossible to move data from a deduplication appliance to other storage.
- Duplicate data extends across numerous storage tiers, including data replicas, archives, and test-and-development copies. Too often, deduplication solutions address only one of those areas. As a result, you end up limiting future opportunities to further reduce your storage consumption.

Candidates for Deduplication

The best candidates for deduplication solutions are midsize or enterprise customers experiencing issues with:

- Exponential growth of data, resulting in out-of-control storage costs.
- Shrinking or inadequate backup windows.
- Longer recovery times, especially for older data not on the primary backup media.
- Cost, risk and complexity of sending tapes to disaster recovery (DR) sites.
- Slow throughput on both backup and archiving systems.
- eDiscovery, compliance and SLA requirements.
- Bottlenecks in expensive LANs and WANs.

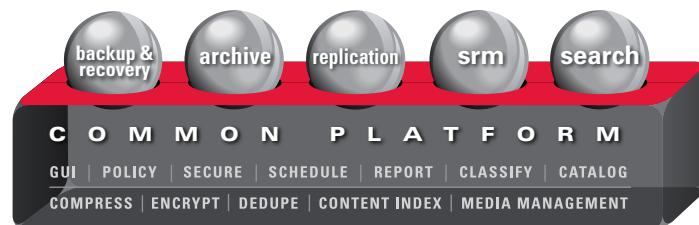
Features to Look For in a Deduplication Solution

When evaluating deduplication solutions, IT decision-makers should look for the following essential features:

- Ability to scale without expensive hardware upgrades.
- More recovery points and with shorter recovery times.
- Point-and-click deduplication management.
- Built-in reporting of deduplication across vendors, data types, sources and platforms.
- Tight integration with all necessary applications to minimize end-user downtime.
- Single solution simplicity for ease of deployment and administration.
- Ability to rapidly and securely recover business-critical data across all locations, applications, storage media and points-in-time.
- D2D2T-optimized for backup performance and reliable data recovery.
- Fast, comprehensive search to aid in recovery.
- Data integrity and security features.
- Built-in Disaster Recovery capabilities.
- Data classification.
- Cost-effective and timely eDiscovery.
- Use of a common technology platform.
- Single point of management.

CommVault® Simpana® Software—A Complete Data Management and Deduplication Solution

CommVault® Simpana® software directly addresses the holistic challenges of data management and integrates all the essential features of a deduplication solution by taking a unique Singular Information Management® approach. This approach employs a single, efficient, scalable platform to perform a complete range of data management functions. Offering an expandable range of application modules, Simpana® software provides multiple backup and recovery options; a flexible IT infrastructure for data protection across physical and virtual servers; and fast, cost-effective disaster recovery to ensure business continuity. In addition to deduplication, Simpana® software includes data archiving, backup and recovery, replication and snapshot management, and data resource management. Used in combination, these capabilities provide unique and compelling benefits for managing data—all from a single unified console.



*CommVault® Simpana® Software's
Singular Information Management® Architecture*

Simpana® software delivers faster and easier data recovery, improved manageability, and reduced costs. Organizations that adopt Simpana® software with its proven, robust, and unified deduplication solution will adapt to change faster and maximize their existing infrastructure. They will be able to offer their developers and customers better application service, responsiveness, and support; and will enable flexible information retention while facilitating compliance and reducing disaster recovery risk. Simpana® software integrates deduplication with other data management functions, it eliminates redundant data across all secondary storage tiers, which makes it possible for you to not only reduce initial storage acquisition costs, but also lengthen the intervals between storage capacity upgrades across both disk and tape. In addition, the vendor neutrality of the Simpana® software deduplication solution—with no hardware lock-in—means easier, simpler administration of the entire data lifecycle. Simpana® software's unique, single architecture keeps compressed data on shared storage resources while delivering fast restores—all managed from a single console.

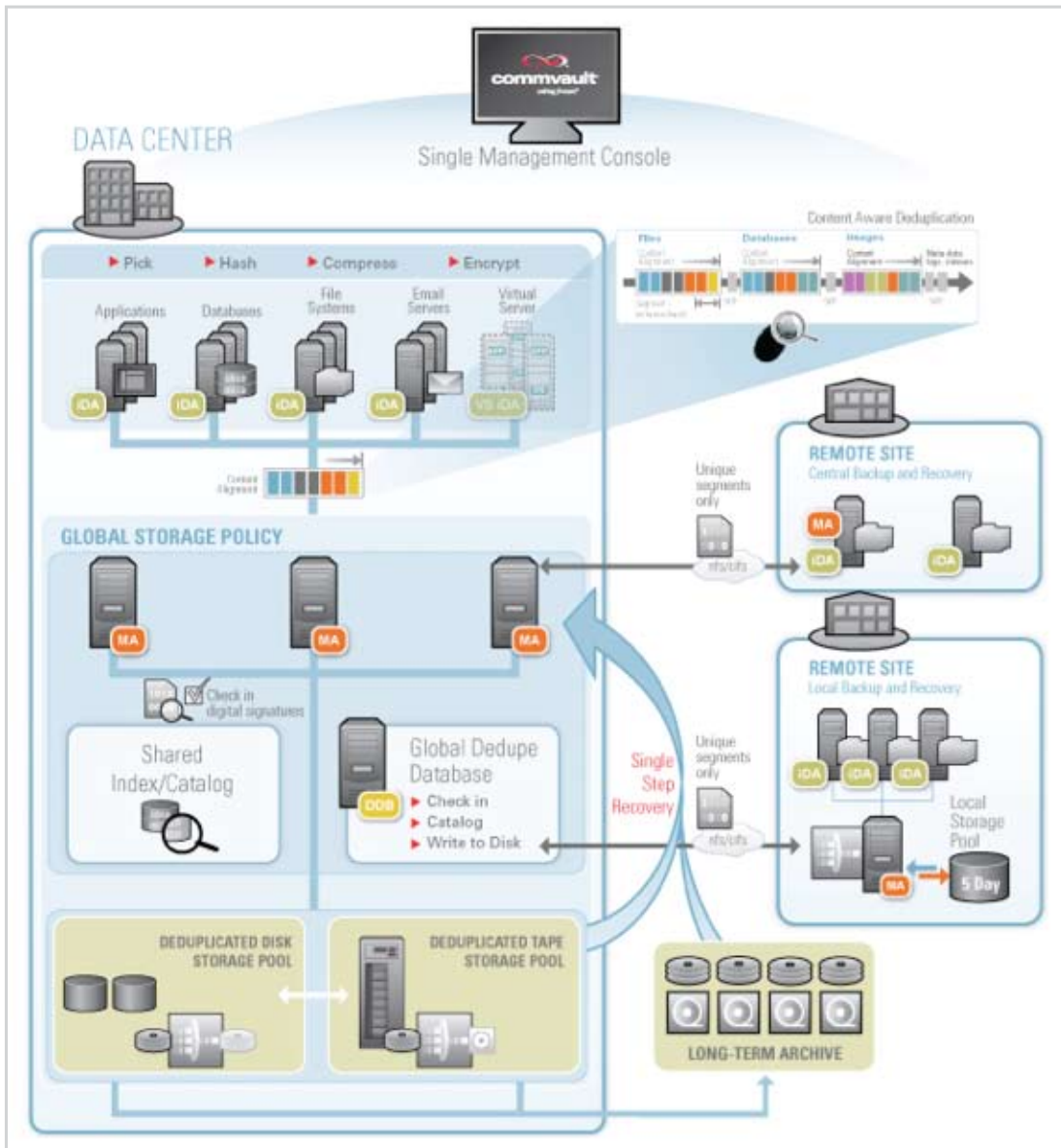


Figure 1: Embedded Global Data Deduplication

Global Data Deduplication for Maximum Storage Efficiency

CommVault extends the power of Simpana® with hardware-independent, content-aware deduplication. CommVault® deduplication software eliminates all duplicate copies of backup and archive data managed within the guidelines of the Global Storage Policy, regardless of source, data type or platform. A combination of content-aware and segment-based deduplication, along with the scalable Deduplication Database delivers better results compared to other block-based approaches. Using the Deduplication Database, the digital signature is compared of each segment against all the other segments of data in the global storage policy, maximizing deduplication across 100% of your backup and archive data. Combining client-based generation of hash signatures with integrated data compression and data encryption capabilities minimizes network bandwidth requirements and maximizes end-to-end network speed and security that goes all the way from clients all the way to disk and tape-based storage tiers.



With Simpana® software you can expand the use of data deduplication across all tiers of storage, including tape, without the need to “rehydrate” the data back into full form. Using our automated, policy-based approach, you can implement a hands-free approach to migrate and manage aging data quickly and efficiently. The migration is invisible to end-users, and ensures that the deduplicated data is brought back seamlessly directly from any storage tier. CommVault® deduplication software also combines the Shared Index/Catalog with the unique Silo Storage to track and retrieve just the data required, without the need for lengthy restores and searching through entire volumes of data.

Deduplication with Simpana is especially helpful for:

- End-to-end elimination of duplicate data across multiple backup cycles, and across backup and archive data stores
→ Greater deduplication ratios.
- Direct data restores from any storage tier and the ability to create and retain dramatically more recovery
→ Faster recovery of data covering longer retention periods.
- Automated policy-based migration/management of long-term data on tape-based backup media
→ Better manageability with optimized storage/tiering of data.

CommVault® Simpana® software integrates deduplication with our broader platform capabilities around Backup and Recovery, Archiving, Replication, SRM and Search to deliver a holistic embedded deduplication approach that maximizes efficiency and minimizes barriers to implementation throughout the enterprise.

Maximizing Deduplication without Sacrificing Performance

1. Content-aware segment-based deduplication across all data within a global Storage Policy that spans all associated clients, applications, agents, jobs, etc.
2. Integration with compression and encryption increases speed and security when writing unique data to disk.
3. Retains dedupe to tape, and eliminates traditional VTL/appliance re-assembly “tax” for direct and selective retrieval of data from any tier (including archived data on tape).

Simpana® Software Benefits

CommVault® Simpana® software offers benefits in the areas of improved performance, improved manageability and increase efficiency.

Improved Performance

- To optimize performance, CommVault advocates a policy of “smart deduplication.” Rather than repeatedly backing up the same data over the network, only to discard it at the backend, CommVault enables a smarter data selection policy that allows only changed objects to be transferred to the backend without impacting recovery. In addition, CommVault maintains a global reference to the existing data segments that are transferred to the backend. In-memory processing ensures redundant data never reaches the disk store. This dramatically improves the scalability of the solution and improves backup performance.
- To deliver rapid restore performance for deduplicated backups, CommVault combines a powerful indexing model with random disk access. During a recovery, Simpana® software reads data segments in the correct order and transfers them to the restore target without reassembly, which delivers maximum restore performance for deduplicated backups.
- By improving backup performance, CommVault enables you to store a larger number of point-in-time backups spanning longer periods of time. Retaining more backups on high-performance disk ensures they can be rapidly recovered. Intact migration of deduplicated data to tape also ensures that restores of older data are even faster and more granular than with alternate approaches. The result is that data recovery has never been so easy, rapid, or powerful. CommVault® deduplication software enables you to cost-effectively meet the most aggressive backup and disaster recovery SLAs, eDiscovery, and compliance requirements.

Improved Manageability: End-to-End Solution Workflow

- Simpana® software's full-featured backup, archive, and data management capabilities work seamlessly with the deduplication feature because they share a common framework, spanning a singular product platform. By moving to Simpana software, you not only get a holistic deduplication solution, you also get a unified view of all your data management processes.
- By delivering deduplication within a complete backup, archival, or disaster recovery (DR) workflow, CommVault simplifies data management. In fact, by integrating deduplication with broader data management functional areas, CommVault simplifies implementation and management, delivering a better ROI in less time. You reap the benefits of improved security at lower costs. For example, utilizing WAN-based remote replication of deduplicated data eliminates the considerable expense and high security concerns of sending truckloads of tapes to DR sites.
- Simpana® software helps assure data integrity and security by providing immediate alerts, which are generated from data integrity checks built directly into the global storage policy model. If any deviation from the original data composition occurs due to accidental alteration of data during transmission or storage, you will be immediately notified. You can also reduce the risks of unintended exposure with client-side deduplication policies that fully support data encryption.

Conclusion

CommVault® Simpana® software is the only end-to-end data management solution that delivers global data deduplication across replication, backup and archiving functions. Its unique, single-platform architecture keeps compressed data on shared storage resources while delivering fast restores, and gives you control over all data management functions from a single console. CommVault's integrated deduplication solution reduces cost and complexity in your IT infrastructure while delivering improved performance, enhanced manageability and increased storage efficiency today and tomorrow.

About CommVault

A singular vision—a belief in a better way to address current and future data management needs—guides CommVault in the development of Singular Information Management® solutions for high-performance data protection, universal availability and simplified management of data on complex storage networks. CommVault's exclusive single-platform architecture gives companies unprecedented control over data growth, costs and risk. CommVault's Simpana® software suite of products was designed to work together seamlessly from the ground up, sharing a single code and common function set, to deliver superlative backup and recovery, archive, replication, search and resource management capabilities. More companies every day join those who have discovered the unparalleled efficiency, performance, reliability, and control only CommVault can offer. Information about CommVault is available at www.commvault.com. CommVault's corporate headquarters is located in Oceanport, New Jersey, in the United States.

For More Information on CommVault solutions for deduplication environments:

www.commvault.com/solutions/deduplication



www.commvault.com ■ 888.746.3849 ■ E-mail: info@commvault.com

CommVault Worldwide Headquarters ■ 2 Crescent Place ■ Oceanport, NJ 07757 ■ 888-746-3849 ■ Fax: 732-870-4525

CommVault Regional Offices: United States ■ Europe ■ Middle East & Africa ■ Asia-Pacific ■ Latin America & Caribbean ■ Canada ■ India ■ Oceania

©1999-2009 CommVault Systems, Inc. All rights reserved. CommVault, CommVault and logo, the "CV" logo, CommVault Systems, Solving Forward, SIM, Singular Information Management, Simpana, CommVault Galaxy, Unified Data Management, QiNetix, Quick Recovery, QR, CommNet, GridStor, Vault Tracker, InnerVault, QuickSnap, QSnap, Recovery Director, CommServe, CommCell and ROMS are trademarks or registered trademarks of CommVault Systems, Inc. All other third party brands, products, service names, trademarks, or registered service marks are the property of and used to identify the products or services of their respective owners. All specifications are subject to change without notice.