

INSIGHT

CommVault Extends Its Data Protection and Information Management Strategy with Simpana 9

Robert Amatruda

IDC OPINION

The release of CommVault's Simpana 9, in late 2010, represents a watershed event for CommVault's core data protection and data management product. Simpana 9 provides its customers with greater flexibility by allowing them to extend their data protection tools beyond their physical environments, safeguarding virtual servers and clients as well. Also, Simpana 9 provides its customers source-side data deduplication that will significantly improve and boost storage asset utilization. Since its inception, Simpana has utilized a common architecture and a single code platform that provides backup and recovery, archive, deduplication, replication, search, and resource management capabilities. We see this as a key differentiator from other prevailing data protection software solutions that typically deliver additional functions via separate software products. Simpana's single data and information management platform provides the underlying code base for functions such as indexing, data movement, cataloging, policy, and information management. CommVault, with its Simpana 9 release, now supports a cloud-based disaster recovery model by connecting to third-party cloud vendors. This capability extends CommVault's cloud-based storage offerings with key service partners such as Amazon S3, Azure, Rackspace, Iron Mountain, and Nirvanix. CommVault continues to evolve its Simpana platform to address the burgeoning data management needs of both physical and virtual environments. We believe CommVault customers have a foundation upon which to deploy a broader set of data and information management functions. New features of CommVault's Simpana 9 software include:

- ☒ Enhanced virtual machine support
- ☒ Source-side data deduplication
- ☒ Expanded support for disk arrays and applications
- ☒ Capacity-based licensing model
- ☒ Migration tools
- ☒ Cloud-based disaster recovery with key service partners

IN THIS INSIGHT

This IDC Insight explores the differentiators of CommVault's Simpana data and information management software and the customer challenges they help address. The focus of this Insight is on the data management and data protection capabilities of Simpana. A more formal document on information management will follow.

SITUATION OVERVIEW

CommVault has made significant investments in its Simpana solution over the past decade to provide a unified data management platform. CommVault's Simpana architecture delivers backup, recovery, archive, replication, storage resource management (SRM), and search capabilities under a single, modular, common code platform. This architecture differentiates Simpana from other prevailing storage software solutions that rely on separate code bases, multiple management interfaces, or additional products that are integrated as part of an acquisition. The benefits of Simpana's common platform and architecture allow customers to scale out their storage environment and take advantage of common management interfaces, integrated search, and fewer processes or products to support.

CommVault's Simpana architecture shares a single set of back-end services, called Common Technology Engine, which ensures granular, release-independent application recovery and services. This capability allows Simpana customers to accelerate new installations and upgrades or add additional data management operations. Moreover, Simpana's native support for email and file archiving under the same framework provides customers greater efficiencies using a single code platform. Simpana's archiving modules built into the common technology platform enable customers to define data management policies to protect active data and archive out-of-date data.

Simpana's common platform and granular management provides its customers with greater simplicity and integration to retrieve items such as messages, contacts, sites, documents, and lists in the context of the data protection process. Additionally, the Simpana platform has flexible search capabilities to capture data sets even if they are embedded inside an application. Simpana provides customers more flexibility, thus lowering capital and operational expenses of deploying separate products. Also, staff training is kept to a minimum, lowering the overall total cost of ownership. Typically, data protection software vendors offer additional products as add-ons for additional fees.

Simpana Architecture Overview

The Simpana architecture is composed of one or more CommServe consoles, MediaAgents, and intelligent data agents (iDataAgents). These components perform the following processes:

- ☒ CommServe communicates with all clients and MediaAgents and coordinates all operations such as backups, restores, copies, and media management within a CommCell.

- ☒ MediaAgent manages the transmission of data between clients and backup media.
- ☒ iDataAgent performs the backup and restore of the clients' data including file systems and supported applications.

The Simpana architecture is managed through a single console — called CommServe. The CommServe acts as a "traffic cop," administering and managing common media agents called intelligent data agents that coordinate and move data from the clients through the CommServe architecture. The CommServe console utilizes a scalable relational SQL 2008 database — that serves as the backup index, storing metadata from the client systems. The database that Simpana software builds and uses is called the CommCell database. The common data movers called media agents reside under control of the CommServe console and CommCell database. The CommServe tells a given client (iDataAgent) to initiate a backup. The data agent crawls the file system and performs backup of the client data. The MediaAgent controls the sending of the data between the client and the backup media chosen (tape, disk, etc.). Once the backup is complete, the client sends the backup metadata to the CommCell, which then commits the metadata to its database.

Split Index Architecture

The Simpana architecture is differentiated because of the use of split indexing. Split indexing provides a more efficient tracking of data elements that are needed for restoration of critical data. The CommServe database stores all metadata, media management, policies, reporting, and licensing information. Furthermore, the CommServe database and media agents have a self-describing index that allows the index to follow the backup job or routine where it physically resides or on any media type. The CommServe database tracks all client or iDataAgent information such as backup jobs through the media sever. The CommServe database can point to the media agent and exact location of the index in the event customers need to restore their data. Thus, all backup and file-level data is kept local. This methodology allows customers to achieve faster performance, more efficient backups, and faster disaster recovery. A Simpana customer does not need to restore an entire monolithic index. By contrast, other prevailing storage software offerings have single, large-scale indexes that need additional management consoles for recovery of data. This adds to complexity, higher latency, and bandwidth limitations.

Also, Simpana allows for automatic protection policies to be added at the end of a backup job to move the data to a disk or tape target. This is achieved by the SQL 2008 database keeping track of the client index regardless of the storage tier. Simpana's architecture supports remote site recovery by passing only the client metadata to the alternate datacenter. A recovery can be achieved, for instance, from physical tapes written in the remote site datacenter and shipped to another datacenter without recovering the master server at the primary site.

High Availability

Simpana's architecture allows customers to have failover and failback for true continuous operations and disaster recovery. Simpana can cluster the CommServe

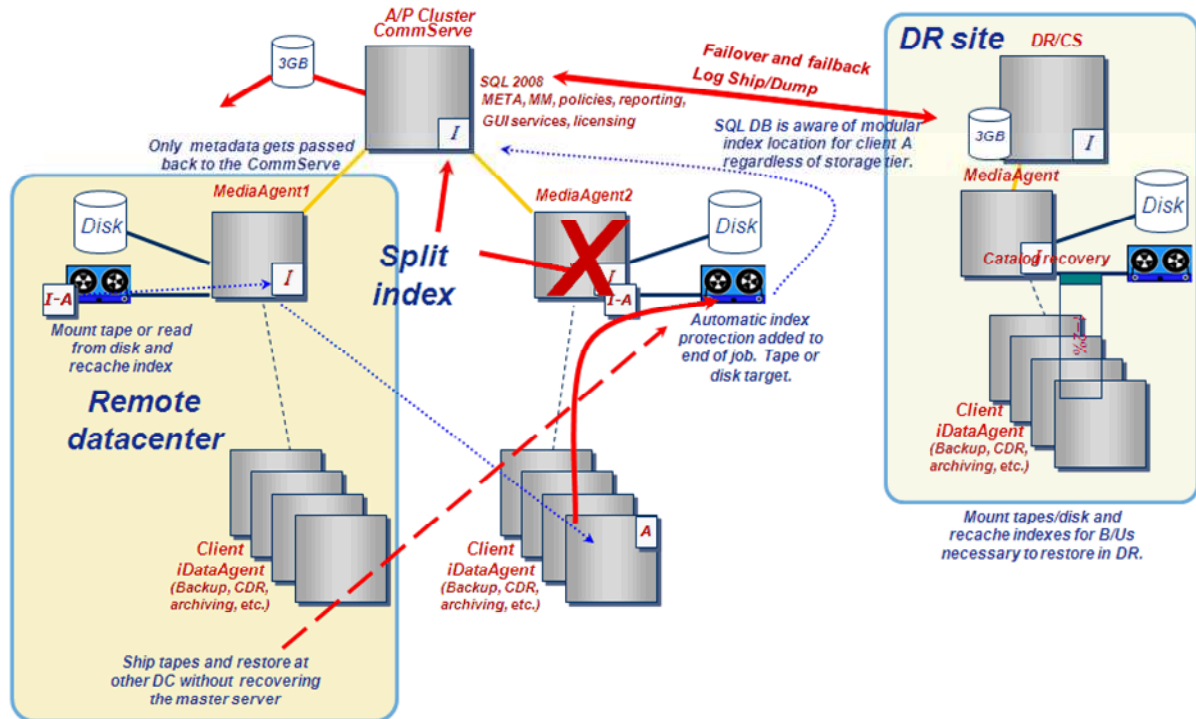
database from the CommServe console to perform failover or failback from a disaster recovery site. Simpana only needs to bring back 1% or 2% of the active catalog data for recovery. Other prevailing solutions need to restore 100% of the catalog for recovery. Simpana can mount disk or tape volumes, recache the index, and perform the recovery.

Multiple Storage Targets

Another benefit of the Simpana architecture is that it allows the iDataAgents to manage data movement and write data to multi-protocol storage targets such as DAS, NAS, SAN, iSCSI, or cloud. Also, the CommServe console and CommCell database capture metadata data, enabling customers to run management reports at any point in the data protection process or during data movement. Using this methodology, Simpana can enable application protection for all backups, archive, SRM, content indexing, or virtual server (see Figure 1).

FIGURE 1

CommVault Simpana Architecture



Source: CommVault and IDC, 2011

Deduplication

In the Simpana 9 release, CommVault delivers its third-generation data deduplication supporting both source and target data reduction while still enabling customers to execute on compression and encryption processes without any negative effect. The customer benefits of data deduplication have been clearly demonstrated by reducing physical storage requirements, providing better network optimization, and reducing costs. CommVault's previous deduplication offerings were based on content-aware, block-based deduplication at the backup server on target storage supporting all backup and archive content. In Simpana 9, the block-level deduplication extends to the source or client server. This allows for deduplicated block data to be checked against the universal set on the server before being stored. In CommVault's Simpana 9 release, customers are experiencing data reduction rates at 90% and better from the source.

The client does this by first scanning its data and then checking the hash marks of that data against that of the media server so only net-new blocks of data are transmitted. Simpana 9 provides its customers an extra measure of flexibility and performance by offering data reduction at the source or client server as well as target or backup server. Providing customers with source-side data deduplication allows greater data transfer efficiencies between clients and backup servers. Additionally, Simpana 9 has increased the scale of the dedupe data store by 50%, meaning it can now handle 64TB+ of data behind a single MediaAgent. It has also added variable block segmentation, providing further storage optimization and reduction. This entire deduplication solution can be deployed on any disk types and is very flexible to the end user's storage architectural design.

Simpana extends deduplication awareness into other data management operations to improve performance and simplify complexity, especially in virtualized environments. Simpana can leverage VMware's vStorage API for Data Protection (VADP) to read and copy data from snapshot copies of data stored on virtual machines. Using VADP, Simpana allows customers to do incremental backups from a snapshot copy of their data. During this process, contents of the virtual machine are cataloged. Thus, Simpana 9 can perform granular restores directly from the snapshot copy of data. Also, multiple copies of data can be created with Simpana's DASHtm Copy (or deduplication accelerated copy) for disaster recovery. Using this method, customers can quickly replicate deduplicated backups on disk, sending only changed data without rehydrating the data or using additional storage resources. Simpana's deduplication methodology allows customers to write and move data to physical tape or copy to cloud storage for long-term retention.

Replication and Availability

In terms of replication and availability, Simpana provides multiple options to its customers. Simpana offers host-based replication that is application aware with what is called the Continuous Data Replication module. The Continuous Data Replication module enables consistent recovery points and continuous backup. This option provides Simpana customers an extra measure of protection when they have stringent recovery time or recovery point objectives. Simpana also integrates with native hardware-based snapshot technologies built into the storage arrays to provide point-in-time recovery with the SnapProtect technology. SnapProtect integrates

seamlessly with applications, virtual environments, and file systems. Simpana eliminates the need for additional scripting needed to integrate array-based snapshots and indexes, thus allowing customers great ease to access and restore snapshot copies of data on the array. Simpana can mount a snapshot on a proxy server to ensure the fidelity of a snapshot as well as browse and execute a completely automated off-host backup to deduplicated disk or tape. Also, snapshot copies can be indexed to provide customers a catalog for viewing data residing on the physical disk array for recovery. This allows for more efficient recovery of data and production copies of data protected.

Virtualization

CommVault continues to address the burgeoning virtualized server and storage customer environments and the data management challenges they encounter. Simpana 9 has expanded its virtual server support to include Citrix Xen as well as Microsoft Hyper-V server and VMware vSphere support. This support enables full guest protection, with granular file-level access for recoveries from a single backup process. Simpana supports backup and recovery of virtual servers — from single files to virtual disks and from virtual systems to physical servers. Integration with VMware's VADP ensures only changed blocks are tracked, which allows for granular recovery.

Simpana 9's Virtual Server Agent (VSA) automatically discovers new non-protected running guests in VMware vCenter. Also, the newly discovered VMware guests will automatically be added to defined protection policies when backups are initiated. Simpana 9 adds the SnapProtect for VSA feature, which is designed to work with integration to array-based snapshot tools to enable rapid creation of consistent recovery copies for hundreds of VMs in minutes without impacting the production VMs while providing for granular file and full VM recovery from the array. This capability also provides the option to create deduplicated secondary copies for disaster recovery or long-term retention, with no impact on physical or virtual machines. Another benefit of Simpana 9 is the ability to do virtual server backups off-host and minimize the impact on production, whether through SnapProtect integration or through streaming API methods. Additionally, Simpana supports migration from physical to virtual environments and virtual to physical environments of different storage platforms.

Expanded Third-Party Storage Array Support

SnapProtect has extended application-aware support for VMware and DB2 in addition to Windows, Linux, Unix, Microsoft Exchange, SQL, Oracle, and SAP. Simpana 9 enables users to extract messages and documents from SnapProtect copies of Microsoft Exchange and SharePoint without requiring the original application. This expedites ediscovery and search requirements while offloading production servers.

Simpana 9 has expanded third-party array support and tiered storage capability. SnapProtect now works with most major storage arrays, with new support for IBM, HP, LSI, and Oracle-Sun arrays. Simpana also supports disk arrays from EMC, NetApp, HDS, and Dell. Simpana 9 allows customers to utilize third-party arrays without requiring any additional scripting. This allows customers to create consistent policies that can span multiple tiers and storage types. For example, a policy could be

set for a "Recovery Tier" to take snapshots every several hours that are retained for a set number of days; a "Protection Tier" could be set to retain a daily backup for 60 days; and a "Compliance Tier" could be stored on disk, tape, or even in the cloud and retained for multiple years. Additionally, Simpana 9 has the ability to recover from any storage tier. This could include a recovery tier using a disk array integrated with CommVault's APIs, a protection tier using disk targets, and an archive tier using disk, tape, or cloud. CommVault's integrated centralized approach minimizes the time associated with manual scripting and aligns a customer's storage tiers according to the customer's needs or requirements. Taken together, these improvements allow customers to more efficiently recover their data from any tier of storage to any target system or platform.

Capacity-Based Licensing

CommVault has also introduced a new capacity-based licensing model for Simpana 9, eliminating charges for most specific agents and options. This new licensing model will allow customers to manage according to SLAs and not SKUs. There are now three basic licensing options: for data protection (per terabyte), for data archiving (per terabyte), and for search/ediscovery (which is based on a per-node/client license). The licensing model is based on the front-end per-terabyte capacity and enables customers to create infinite copies without increasing their costs.

Fast Pass Migration Tools (NetBackup or TSM)

Simpana 9 simplifies the migration process with Fast Pass, which extracts metadata and job results — such as discovering backup policies, client configuration settings, job details/history, schedule policy, and job inventory from either the NetBackup (6.0, 6.5, 7.0) or the TSM (5.4–6.1) master server and imports it all into a new CommVault client configuration, which then automatically performs the agent installation process. The old system can be left running for historical recoveries only.

Other Simpana 9 Innovations

- ☒ **Enhanced NetApp replication integration.** NetApp storage customers using SnapVault and SnapMirror with SnapProtect can take local application-consistent snapshots and replicate the data sets to another NetApp array only moving changed data.
- ☒ **Cloud storage support.** Simpana 9 has added support for cloud storage platforms from Dell (DX) and HDS (HCP), in addition to existing support for cloud platforms from Amazon, Azure, EMC (Atmos), Iron Mountain, Meseo, Nirvanix, and Rackspace.
- ☒ **Enhanced SRM reporting and SaaS remote services.** Simpana 9 has improved storage resource management featuring integrated reporting that automatically reassesses daily usage and provides threshold warnings along with trending analysis.
- ☒ **Enhanced application support.** Simpana 9 expands application support for Microsoft Exchange 2010 and DB2 to ensure rapid recovery and improved SLAs.

- ☒ **SAP archiving enhancement.** For SAP customers, CommVault offers new integration capabilities including policy-based archiving for improved database and application performance.
- ☒ **Integrated management of array-based snapshots technology.** Simpana 9 offers integrated management of array-based snapshots technology. This feature, called SnapProtect, manages, controls, and retains data across a multitude of array-based snapshot vendors, applications, and operating systems.

Simpana supports file archive for multiple platforms with the same media agent and supports tiered storage options such as disk, tape, optical, and CAS systems. The Simpana email archiving solution can be content indexed and supports multiple media types for low-cost tiered storage. This solution can navigate, find, and select messages to retrieve with point-and-click browse of the archive using an Exchange-like folder tree. In addition, Simpana also supports archiving for Exchange, Lotus Notes, and SharePoint.

FUTURE OUTLOOK

CommVault's theme of "modernized data protection" will resonate well with CommVault's current installed Simpana customers as well as those customers looking to bridge their data protection solutions from physical to virtual to cloud. We believe Simpana 9's enhancements fit very well into CommVault's cloud offering by utilizing content indexing, data reduction, tiering of storage resources, and improved management. These enhancements in the version 9 release will enable Simpana customers to migrate and archive data to lower-cost cloud-based storage. CommVault has set the foundation for its customers to protect their data from core to cloud.

CommVault has had very good success in the marketplace, growing its installed base to well over 14,000 licensed customers. In addition, CommVault has penetrated large global OEMs, such as Bull, Dell, HDS, Hitachi Limited, NetApp, and Fujitsu. In addition, CommVault has numerous two-tier distribution agreements in place. We believe CommVault's Simpana 9 release will position the company more competitively against well-entrenched vendors in the data protection software market, with the extension of its data management infrastructure along with expanded VM support, source deduplication, and more integrated replication management. We expect CommVault, with its Simpana 9 offering, will continue to enjoy attractive growth rates and aggressively target its competitors in the marketplace. CommVault's singular focus on data and information management positions CommVault well for future growth in the data protection software market.

CommVault Simpana's differentiators and strengths include:

- ☒ CommVault's Simpana software is delivered on a common platform with data protection, archive, replication, and search capabilities and supports all open system platforms. The Simpana software integrates and embeds block-level source deduplication throughout a firm's data management infrastructure.

- ☒ CommVault has a singular focus on data and information management and is unencumbered by acquisitions and product integration or protection of a legacy installed base.
- ☒ CommVault is an independent software company and not a system vendor, which many customers will find attractive as they don't want to feel locked-in to a single vendor solution.
- ☒ CommVault Simpana has attractive growth rates and is aggressively targeting the enterprise market segment, remote-office and back-office protection, and cloud.

Copyright Notice

This IDC research document was published as part of an IDC continuous intelligence service, providing written research, analyst interactions, telebriefings, and conferences. Visit www.idc.com to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit www.idc.com/offices. Please contact the IDC Hotline at 800.343.4952, ext. 7988 (or +1.508.988.7988) or sales@idc.com for information on applying the price of this document toward the purchase of an IDC service or for information on additional copies or Web rights.

Copyright 2011 IDC. Reproduction is forbidden unless authorized. All rights reserved.