commvault® *solving forward®*

Advanced Deduplication with CommVault® Simpana® 10

Highlights:
Simpana® software deduplication benefits extend further across the enterprise and the cloud with 4th Generation Deduplication technology.

1. **Enterprise-class Parallel Deduplication** provides increased throughput, resiliency, and scalability by distributing workloads across multiple nodes.

2. **Global Deduplication** enables the deduplication of multiple datasets across different locations with different storage systems and storage policies.

3. **Source side deduplication** eliminates redundant backup and archive data at the client, before using expensive network resources to transfer data.

4. **DASH Copy and DASH Full** capabilities dramatically reduce the time required to copy deduplicated data to other locations and deliver individual retention policies for Simpana 10 backup and archive copies.

5. **Enhanced performance and scale with Solid State Drives (SSDs) for Deduplication database**

4th Generation Deduplication

Deduplication has rapidly become a standard requirement for many IT architects looking to reduce storage costs. Building on its innovation in this area, the fourth generation of CommVault® deduplication technology scales beyond single nodes by introducing Parallel Deduplication, which allows you to multiply your capacity, throughput, and resiliency by consolidating multiple deduplication nodes under a single global reference pool.

Global Source and Target Side Deduplication Built-In

- **Reduction of Storage Consumption** - Reduce secondary storage needs by up to 50% with embedded deduplication.

- **Reduce or Eliminate Network Constraints** –reduce network traffic by up to 90% with source-side deduplication resulting a deferral or cancellation of costly network upgrades.

- **Better Protection Strategy:** Implement smarter protection strategies like incremental forever for supported data types, limiting the impact of backups on production server while leveraging DASH Full technology to enable wide variety of recovery options, including single pass full restore from incremental backup.

- **Global Deduplication** – Deduplicate across multiple data sets, and multiple storage policies. This provides maximum flexibility while minimizing storage acquisition costs.

- **Efficient Copy Management** – Create unique deduplicated copies of data with independent retention policies, and store onsite, at remote sites, or even in the cloud without having to rehydrate the data using Deduplication Accelerated Streaming Hash (DASH Copy).

- **Consolidate Remote Office Data Protection** – Reduced complexity and streamline operations by leveraging source-side deduplication and DASH Copy capabilities to eliminate remote office backup silos.

Introducing Parallel Deduplication

- **Enterprise Class Scalability** – Protect large scale deployments using a "grid" based architecture that combines multiple deduplication nodes into one master set. This enables linear scale out on store size, which is key for Enterprise or Global consolidation.

- **Faster Backups and Recovery** – Increase throughput and concurrency by distributing workloads across multiple nodes delivering linear scalability at a rate of 4.5 TB/hr per deduplication node.

- **Resilient Backups** – Enable failover protection with Parallel Deduplication by automatically redirecting deduplication processing in the event one node is temporarily inactive.

Scalable Deduplication from Very Small to Very Large

The breadth of CommVault's deduplication technology allows customers tremendous flexibility, ranging from a single node with a small deduplication database up to a very large multi-node deduplication database that uses Parallel Deduplication for enterprise-class scalability. Performance and scalability can be further enhanced by using SSD drives, even in single node configurations.

Node Type	Suggested DDB Vol Size	Estimated Backend Size*	Estimated Front End Size **	Sample Configs
XL	800GB – 1TB	Up to 120TBs	Up to 60TBs	Fusion IO Drive2 785GB PCI Express 2.0x4 MLC SSD; **OR** 4xSSDs (NAND Flash) w/ 600GB w/ 3 spindle RAID 5 w/ 1 HS config
Lar	800GB – 1TB	60 – 90TBs	Up to 40TBs	8 spindles, 15K RPM, SAS/FC, 300+GB ea, RAID 10 config
Med	600GB	30 – 60TBs	20 – 30TBs	6 spindles, 15K RPM, SAS/FC, 300+GB ea, RAID 10 config
Sm	400GB	Up to 30TB	Up to 20TB	4 spindles, 15K RPM, SAS/FC, 300+GB ea, RAID 10 config

commvault *solving forward*

KEY CAPABILITIES AND BENEFITS

Partitioned Deduplication	New grid based architecture enables large enterprise class scalability. Multiple deduplication nodes can now be combined into one master node providing near linear scale out on store size.
Resilient Backups	With new Parallel Deduplication, failover protection is enabled automatically. When one node goes down, the remaining nodes take over the deduplication processes.
Increased Throughput	Parallel deduplication provides unmatched performance, linearly scaling throughput up to 4.5TB/hr as workloads are distributed across multiple nodes (for two nodes).
Source Side and Target Side Deduplication	Protect remote offices or data centers with better network efficiency. Deduplication can be enabled at both the source (client server) and the target (backup server). Source side deduplication minimizes the time needed to protect remote data and minimizes WAN requirements. Target side deduplication provides a second filter that maximizes deduplication across clients before data is finally stored on disk.
Deduplicate and Recover Directly From Multiple Storage Tiers – from a single console	CommVault Simpana software enables users to deduplicate and recover data seamlessly from multiple types of storage, including disk, tape, and cloud based platforms – from a single console.
Modular, Scalable and Compatible With Open Storage Platforms	CommVault's open approach to deduplication eliminates the lock-in of appliance based approaches. Any open disk platform can be used if it meets the requirements of your data protection strategy. Increasing capacity is as simple as adding more storage to your deduplication data pool, regardless of the original storage arrays or vendors.
Deduplicate Data Globally	Data deduplication can be extended to cover multiple data sets, and multiple storage policies. This provides users with maximum flexibility for data management, while minimizing storage acquisition costs.
Efficient Copy Management and Flexible Retention with DASH Copy	Create deduplicated copies of data with independent retention policies on disk or tape, and store onsite, at remote sites, or even in the cloud without having to rehydrate the data using Deduplicated Accelerated Streaming Hash (DASH Copy).
Deduplication Reporting and Analytics	New Parallel Deduplication functionality allows users to use the Deduplication Summary to monitor Deduplication Savings across multiple Storage Policy Copies in the CommCell and monitor the amount of data written on the disk after deduplication.
Smarter Protection Policies	Use smarter techniques instead of brute force data movement enabling all of the recovery capabilities with limited impact on production servers.

commvault *solving forward*