FUTUREWEI
Technologies

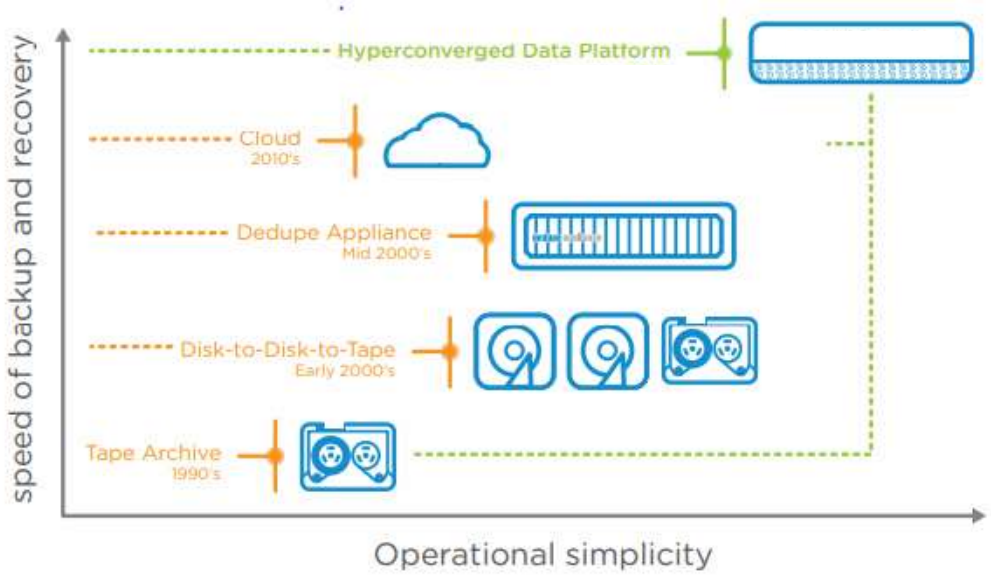
Backup & Recovery Competitive Analysis

Futurewei Intelligent Data Lab

3/2/2021

Introduction

Evolution of Backup & Recovery



© [Cohesity®_Data_Protection_White_Paper.pdf]

Challenges:

- Speed of backup & recovery measured against delivery on SLAs
- Complexity of operational maintenance caused by multiple siloed hardware and software, heterogeneous management tools and islands of storage subsystems
- Total cost of ownership

FUTUREWEI Technologies

Key Requirements

- Simplify backup & recovery
- Speed up backup, while eliminate the application impact
- Rapid recovery, reduce RPO, and ensure business continuity
- Off-site data protection
- Easily on-ramp to hybrid and multi-cloud
- Reduce data and storage footprint (duplication and compression)
- Reduce risk / Ransomware protection
- Long time data retention for compliance and regulatory requirements
- Reduce cost & save space

Gartner 2020 MQ for Backup & Recovery Solutions

Source: Gartner (July 2020)

Veeam® : Leaders for 4th time highest for “ability to execute”

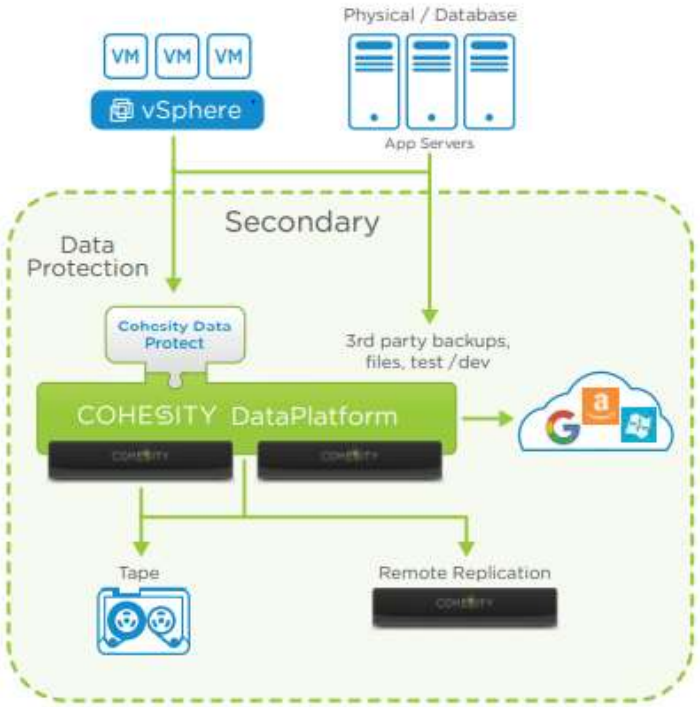
Dell® : Leader over 20 years

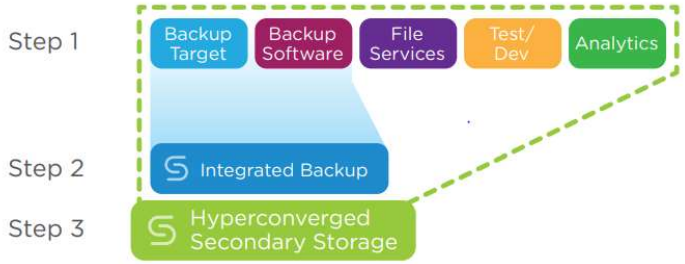
Cohesity® : Visionarie -> Leader

FUTUREWEI Technologies

Cohesity®

Hyper Converged Secondary Storage





- DataProtect® provides tight integration with VMware vCenter
- Support 3rd party backup application

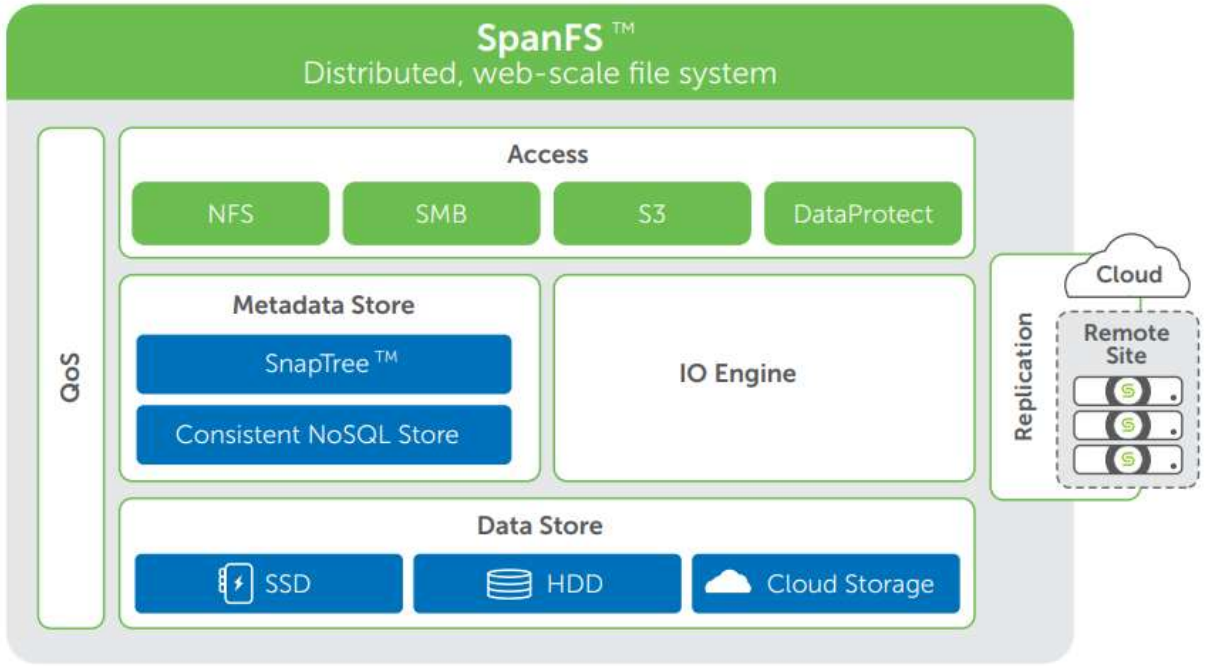
- Cloud archiving the older local snapshot
- Cloud tier between Cohesity® cluster and the cloud
- Cloud replication

- Long-term data retention to external tape and public cloud

© *Cohesity®_Data_Protection_White_Paper.pdf*

FUTUREWEI Technologies

SpanFS™



© [Cohesity® SpanFS and SnapTree]

The data is distributed across all the nodes in the cluster. Volumes are accessed through a virtual IP mount point, and user access and IO are distributed across the nodes using the virtual IP address.

The IO Engine automatically detects whether the workload is sequential or random and directs IO to the appropriate data path and media tier (HDD or SSD).

SnapTree™ (B+ tree based):

- Distributes the tree across nodes
- Supports the creation of instantaneous clones and snaps
- Garbage collects unreferenced nodes in the background using Map-Reduce
- Stores only one value per leaf node
- Supports a variable fan-out factor

FUTUREWEI Technologies

DataPlatform

- Scale-out (3 nodes at least) , hyperconverged, pay-as-you-grow model
- **Distributed file system** inspired by google file system. Share-nothing.
- Unified platform for end-to-end data protection (backup, restore, replication, DR and failover/failback orchestration).
- zero-cost, highly scalable **snapshots** and clones for fast recovery point. Cohesity® SnapTree™ manages data copy with B+ tree.
- **Index engine** that indexes all the VMs being backed up, the files within the VM and associated metadata.
- Parallel data **ingest engine** handles data placement and adaptive data throttling over the workloads at the vCenter or datastore level.
- Long-term data retention to external tape and public cloud.
- Global variable-length **deduplication**. Allow customer to decide in-line or post-process.
- **Compression** of deduped blocks.
- **Replication**: one-to-one, one-to-many, many-to-one, and many-to-many models.
- CloudTier™, CloudArchive™, and CloudReplicate™
- Encryption engine: data at rest and in transit
- Role-based access control

Performance Improvements

- Fast backups with parallel ingest
- Workload-optimized ingest: Tier-Optimized Write Scheme (TOWS), disk for sequential I/O and SSD for random I/O
- Incremental snapshots with 15-minute RPOs
- All incremental available instantly
- Instant restore to any point-in-time copy

Advantages

Simplicity and Automation

- Converged Backup and Recovery
- Unified User Experience: one UI to manage and provision from the core, to edge, and public cloud.
- Policy-based Automation automates operations and meet your business SLAs.

Globally Search and Instantly Recover

- Rapid Recovery Points reduce recovery time with fully hydrated snaps, without any performance impact.
- Fast Global Search across all workloads and locations from a single UI.
- Instant Mass Restore recovers VMs and files, plus application objects instantly. Speed up large file recovery with MegaFile.

Enterprise-class Performance and Security

- Ransomware Protection: Immutable snapshots with AES-256 encryption, multifactor authentication, and FIPS certification.
- Data Resiliency at Scale
- Industry leading space efficiency with global variable length deduplication across workloads and protocols.

© *Backup and Recovery Software | DataProtect® | Cohesity®*

Delivered the Way You Want

- Backup as a Service
- Software Subscription: Run in public clouds and on-prem on certified platforms.
- Service Provider Managed: Let service provider manage on your behalf.

FUTUREWEI Technologies

Disadvantages

- Cannot mix and match generations of equipment in the same enclosure
- No file-level restore, or a single VM restore
- Replication performance is not good (may be improved in newest releases)
- OS backup compatibility issues
- Management: need more analysis
- Security features: need more analysis
- Multi-tenancy has limited scope – not across all platforms
- Documentation – support matrix is not up to date
- Need to support more products (Sharepoint, etc.)
- Need to support more storage vendors

C6000 Series

	Model			
Description	**C6025**	**C6035**	**C6045**	**C6055**
Hard disk capacity	96 TB	144 TB	168 TB	192 TB
PCI-e based flash storage	6.4 TB	6.4 TB	6.4 TB	6.4 TB
Memory	128 GB	128 GB	128 GB	128 GB
CPU	2 x Intel Xeon (2.3 GHz, 12 Core)	2 x Intel Xeon (2.3 GHz, 12 Core)	2 x Intel Xeon (2.3 GHz, 12 Core)	2 x Intel Xeon (2.3 GHz, 12 Core)
On-board network connectivity	4 x 10 Gb or 2 x 25 Gb or 2 x 40 Gb and 1x IPMI	4 x 10 Gb or 2 x 25 Gb or 2 x 40 Gb and 1x IPMI	4 x 10 Gb or 2 x 25 Gb or 2 x 40 Gb and 1x IPMI	4 x 10 Gb or 2 x 25 Gb or 2 x 40 Gb and 1 x IPMI
Dimensions	W x H x D (in): 17.6 x 1.7 x 34.7 W x H x D (mm): 448 x 43.2 x 881			
Weight	70 lbs/30 KG			
Operating environment	Cohesity Helios multicloud data platform			
Storage protocol support	NFSv3, SMB 2.1, SMB 3.0, S3 REST API			
Data protection integrations	VMware VADP, Oracle RMAN, Microsoft SQL, Windows, Linux, Pure Storage, NAS, Nutanix AHV, Microsoft Hyper-V, NetApp			
Data resiliency	Erasure Coding and RF/2			
Cloud archive	Google Nearline, Microsoft Azure, Amazon S3 & Glacier, any S3 and NFS compatible storage			
Encryption	AES256			
Power supply	2x Redundant (1+1) PSU with PMBUS			

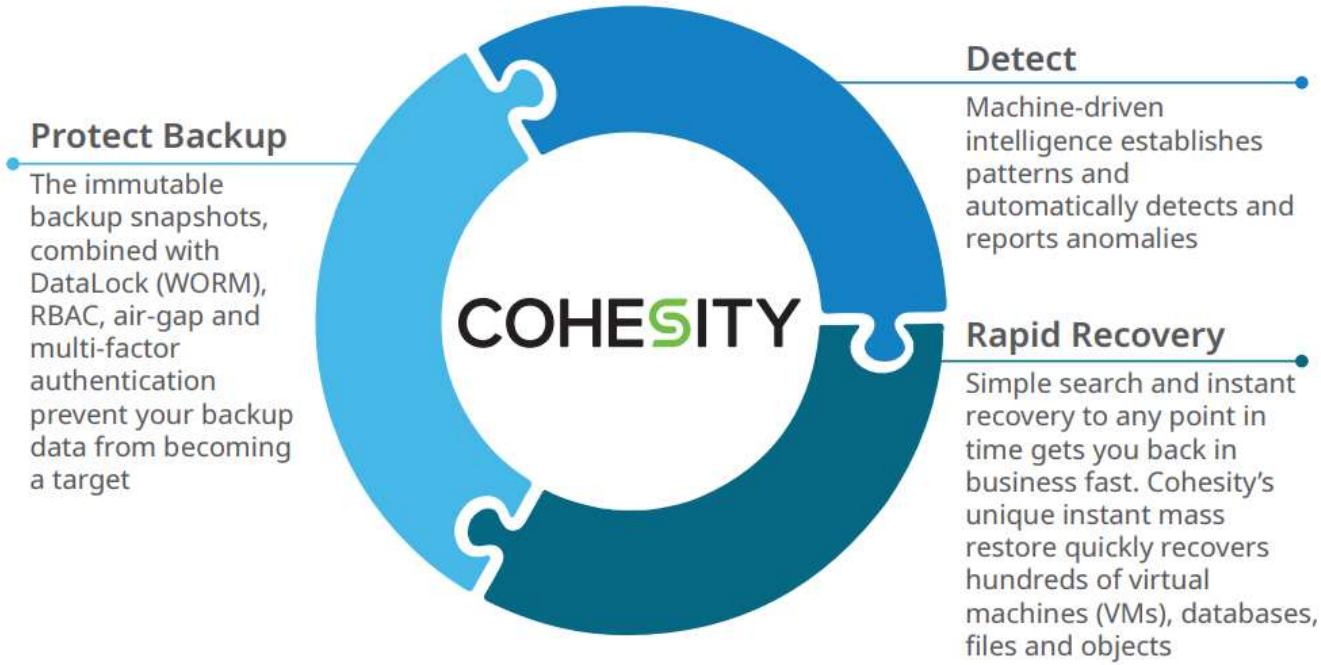
NAS scale out limits:

- Number files per directory: 3.2 million
- Max path length:131092
- Max file size: 70 TiB
- Max number of snapshots: 1 million

© *Architecture Matters (Cohesity®.com)*

FUTUREWEI Technologies

Ransomware Recovery



© Simplify Long-Term Data and Application Retention and Archival (Cohesity®.com)

EMC®
PowerProtect®

DD Series

Features:
Global Compression™,
Data Invulnerability Architecture, including inline verification and integrated dual disk parity RAID 6, snapshots, telnet, FTP, SSH, email alerts, scheduled capacity reclamation, Ethernet failover and aggregation, Link Aggregation Control Protocol (LACP), VLAN tagging, IP aliasing, DD Boost, DD Encryption, DD Extended Retention, DD Retention Lock, DD Virtual Tape Library (VTL) (for open systems and IBMi operating environments).
Available add-ons include DD Boost, Cloud Tier for long-term retention, Cloud Disaster Recovery, and DD Replicator™.

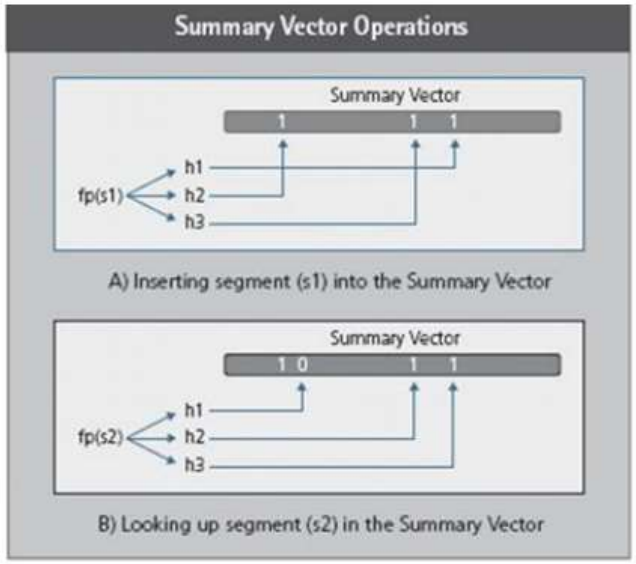
Data Management:
NFS v3 over TCP, CIFS and DD Boost™over 1GbE or 10GbE or Fibre Channel, tape library emulation (VTL) over Fibre Channel, and NDMP Tape Server.

Software-defined appliance with PowerProtect® DD Virtual Edition

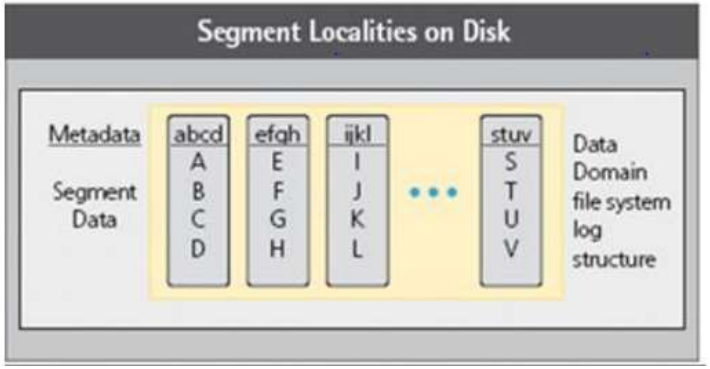
© [DellEMC® PowerProtect® DD Series Appliances with Commvault® Backup and Recovery : Configuration Guide (Dell®technologies.com)]

FUTUREWEI Technologies

Stream-Informed Segment Layout (SISL™)



The summary vector is an im-memory data structure to quickly identify new, unique segments. Based on a Bloom filter, the summary vector is a bit array in RAM initialized to 0s. When a new segment is stored, a few bit locations in the array are set to 1s. The locations are chosen based on the fingerprint of the segment.



Segment localities: Data Domain stores neighboring segments together as a sequence of segment units called segment localities, along with fingerprints and other metadata, into containers and appended to the log of containers. The file system is a log structured system, which is the log of containers storing localities.

© *Dell® EMC® Data Domain SISL ™ Scaling Architecture (Dell®technologies.com)*

FUTUREWEI Technologies

Features

- **Data integrity**—The DD OS Data **Invulnerability Architecture** provides:
  - Inline end-to-end write/read verification (checksum)
  - Append only write, log-structured file system
  - Validation after a backup is complete
  - Background continuous fault detection and self healing
  - file system recovery with self-describing metadata rebuild.
  - Raid 6 protection
- **Data Deduplication**—The file system deduplicates data inline in CPU and memory. Variable-length segmentation.
- **Compression** — lz/gz/gzfast. Support hardware accelerated, gzfast compression algorithm
- Restore operations—File restore operations create little or no contention with backup or other restore operations.
- **Multipath and load balancing**—When multiple paths are present, the system automatically balances the backup load between the available paths.
- **High availability**—The High Availability (HA) feature lets you configure two protection systems as an Active-Standby pair, providing redundancy in the event of a system failure.
- Random I/O optimizations provide improved performance
- **Scalability**: shelves can be added while the system is running providing additional scale without disruption. Data Domain Cloud Tier and extended retention options.

Features (cont)

- **Stream-Informed Segment Layout (SISL™)** scaling architecture enables Data Domain systems to perform 99% of the deduplication processing in CPU and RAM, which gives it fantastic performance even with inefficient protocols like CIFS and NFS. It stores related segments and fingerprints together so large groups can be read at once.
- **DD Boost™**software distributes parts of the deduplication process to the backup server(s) or application client(s), leaving the Data Domain system to focus its energy on determining what is unique and writing the new data to disk. With DD Boost, only the unique data has to travel from the backup server or client to the Data Domain system. DD Boost™also gives the backup application control over replication. DD Boost™also enables automatic load balancing of the backup workload across all the available paths to maximize performance and efficiency. In addition, DD Boost™provides automatic path failover.
- **DD Replicator™**software are designed to improve disaster recovery supporting one-to-one, one-to-many, many-to-one, many-to-many and cascaded replication topologies.
- DD Retention
- Secure **multi-tenancy** capabilities enable secure data isolation, management and reporting by internal business units or departments, or individual customers.
- **ProtectPoint** uses time point based snapshot.

Advantages

- Optimized write performance (Up to 41 TB/hr)
- With DD Boost™, max throughput up to 94 TB/hr
- SISL™ enables systems to perform 99% of the deduplication processing in CPU and RAM
- Log structured file system
- DD Boost™ provides source side deduplication, reducing ingest bandwidth
- VTL interface enables non-disruptive roll-out into existing backup infrastructure
- Invulnerability Architecture provides full data integrity protection

Disadvantages

- Scale-up architecture
- HA is active-standby, need up to 10 minutes to fail over
- OS upgrades are slow and disruptive
- Deduplication is limited to one controller
- DD series is only a backup target. It cannot schedule, manage, catalog, initiate restores or provide visibility into backups
- Slow to restore, compared to fast backup
- Need at least 20% free space to work properly
- Offloading to cloud storage needs expensive license
- TCO can be high: the larger the appliance, the more you pay per TB

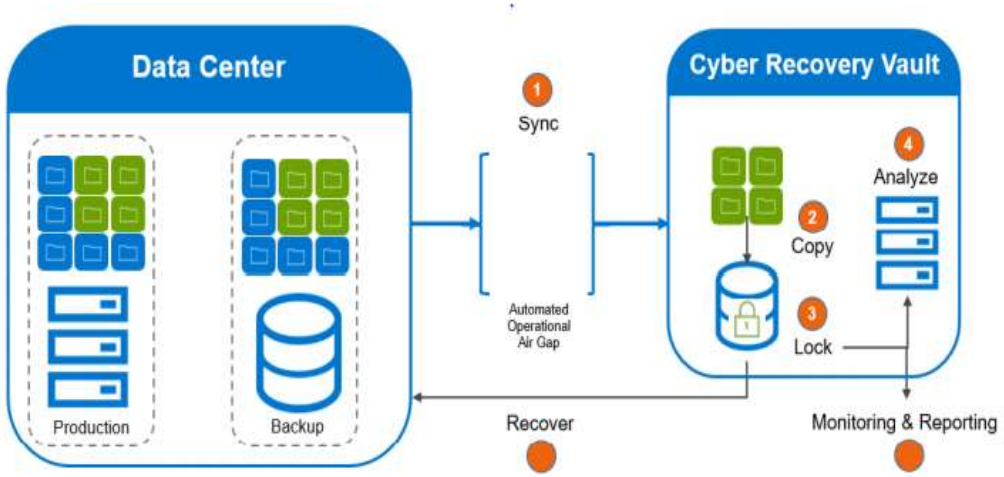
DD Series Spec

	DD3300	DD6900	DD9400	DD9900
Max Throughput	Up to 4.2 TB/hr	Up to 15 TB/hr	Up to 26 TB/hr	Up to 41 TB/hr
Max Throughput (DD Boost)	Up to 7.0 TB/hr	Up to 33 TB/hr	Up to 57 TB/hr	Up to 94 TB/hr
Logical Capacity[1]	Up to 1.6PB	Up to 18.7PB	Up to 49.9PB	Up to 97.5PB
Logical Capacity with Cloud Tier	Up to 4.8PB	Up to 56.1PB	Up to 149.8PB	Up to 228PB
Usable Capacity	4TB – 32TB	24TB – 288TB	192TB – 768TB	576TB – 1.5PB
Usable Capacity with Cloud Tier	Up to 96TB	Up to 864TB	Up to 2.3PB	Up to 3.5PB
ES40 Shelf	N/A	4TB 7.2K SAS	8TB 7.2K SAS[3]	8TB 7.2K SAS[3]
DS60 Shelf	N/A	4TB 7.2K SAS[3]	8TB 7.2K SAS	8TB 7.2K SAS
FS25 Shelf	N/A	3.8TB SSD[2]	3.8TB SSD[2]	3.8TB SSD[2]

File system limits:
- 1 billion files
- 2 billion inodes (files and directories)
- Maximum length of a path name is 1023
- During HA failover access to files may be interrupted for up to 10 minutes

© *h12927-Dell®EMC®-PowerProtect®-dd-ss.pdf (Dell®technologies.com)*

Cyber Recovery®



Proven and Modern Protection for Critical Data from Ransomware and Destructive Cyber Attacks

Automated Data Copy and Air Gap
Create unchangeable data copies in a secure digital vault and processes that create an operational air gap between the production / backup environment and the vault.

Intelligent Analytics and Tools
Machine learning and full-content indexing with powerful analytics within the safety of the vault. Automated integrity checks to determine whether data has been impacted by malware and tools to support remediation if needed.

Recovery and Remediation
Workflows and tools to perform recovery after an incident using dynamic restore processes and your existing DR procedures.

© *isolated-recovery-solution-overview.pdf (Dell®technologies.com)*

FUTUREWEI Technologies

Comparisons

Spec

	DD 9900	Cohesity®	OceanProtect®
Max throughput	• Up to 41 TB/hour	• Info N/A	• Up to 94 TB/hour
Capacity	• Logical up to 97.5 PB • Usable up to 1.5 PB	• 192 TB per appliance (C6000)	• Usable up to 9.2 PB
Capacity with cloud tier	• Logical up to 228 PB • Usable up to 3.5 PB	• Cloud backup • Cloud archival	• Cloud backup • Cloud archival
Workloads	• VMware and open hypervisors, including Nutanix, Citrix, Red Hat, KVM, and others • Kubernetes containers • Oracle • Microsoft SQL and Exchange • SAP HANA • Windows and Linux Filesystems • Cloud: EMC® ECS, AWS, Azure, Virtustream	• Hypervisors: VMware vSphere, Microsoft Hyper-V, Nutanix AHV and RHeV • Kubernetes-based data and application state • Cloud: AWS EC2, Amazon RDS, Azure VM, and Google Compute • Physical: Windows, Linux (RHEL, CentOS, OEL, Debian, Ubuntu), AIX (7.x), and Solaris • Databases: Oracle RAC and SQL • Distributed Databases (NoSQL and Hadoop): MongoDB, CouchbaseDB, Apache, Cassandra, Cloudera and Hortonworks • Bare Machine Recovery: Windows and Linux • Applications: MS Exchange, MSFT Active Directory, Microsoft 365 , SalesForce (SFDC), and SAP HANA • Primary storage: Pure FlashArray, HPE Nimble and Cisco Hyperflex NAS: Pure FlashBlade, NetApp, Isilon*, IBM GPFS, Google EFS, Elastifile, and generic solutions	• Info N/A

Features

	DD	Cohesity®	OceanProtect®
Scalability	• Only scale up with disks/shelves	• Scale out, share-nothing architecture	• Scale out, share-nothing architecture
File system	• DDFS, log-structured file system	• SpanFS™, distributed file system (similar to GFS)	• Distributed file system
Interfaces	• NFS v3 over TCP, CIFS and DD Boost™over 1GbE or 10GbE or Fibre Channel, DD VTL over FC, NDMP Tape Server	• NFS, SMB, S3	• NFS, CIFS, SMB • DPC supports POSIX, and MPI-IO protocols
deduplication	• client-side deduplication (DD Boost™) • Inline deduplication (SISL™) • Typical ratio 20:1	• Global variable-length data deduplication • Inline or post-process with user config	• Global variable-length deduplication • Inline + post-process • Similarity-based deduplication
compression	• lz/gz/gzfast • Hardware acceleration	• Compression of deduped blocks	• proprietary compression algorithm to compress the data that is difficult to compress • Inline and post-process compression
Reliability	• Raid 6 • Inline Verification	• Multi copy or EC between nodes	• Multi copy and EC
HA	• Active-standby • Failover takes up to 10 minutes	• Active-active, distributed	• Active-active, distributed, support multiple controller failure
Snapshot	• Time point based	• Distributed snaps and clones with no performance impact • 15-minutes RPO	• Time point based • Second-level RPO
Replication	• Asynchronously replicates compressed, deduplicated data over a WAN • Support full system mirroring, bi-directional, many-to-one, one to-many, and cascaded	• Unlimited geo replication. • Support one-to-one, one-to-many, many-to-one, and many-to-many models	• Sync and async replication • 3DC • Support one-to-one, one-to-many, many-to-one, and many-to-many models

Features (cont)

	DD	Cohesity®	OceanProtect®
Multi-tenancy	• Supported with Mtrees • RBAC	• multitenancy with strong QoS support, data isolation, separate encryption keys, and RBAC	• Multitenancy, QoS, RBAC
Tiering	• DD Extended Retention: internal tiering within active tier and a retention tier • Cloud tier	• Automatic tiering of data between SSD, HDD, and cloud storage	• Support tiering to OceanStor Pacific series
Security	• DD Encryption, AES-256, system level • Data at rest and in transit • DD Retention Lock	• Software-defined AES-256 • Data at rest and in transit	• Software-defined AES-256 • Data at rest and in transit
Cloud	• Cloud Tier for long-term retention • Cloud Disaster Recovery • EMC® ECS, AWS, Azure, Virtustream	• For archival, tiering and replication • AWS, Microsoft Azure, Google Cloud	• Cloud backup • Cloud archival
Add-ons	• DD Boost™ • Stream-Informed Segment Layout (SISL™)		
Archival	• Cloud, archival tier, archival device	• Cloud, tape, S3 and NFS devices	• Cloud, blue ray
3rd-party backup software support	• Commvault, Veritas NetBackup, Veeam	• Commvault, Veritas NetBackup, Veeam	
Deliveries	• DDVE® sds • DD series	• Backup as a Service • Software Subscription: Run in public clouds and on-prem on certified platforms. • Service Provider Managed	• OceanProtect® X/A

Secure Recovery

	DD	Cohesity®	OceanProtect®
Product Name	• **Cyber Recovery®**	• **Ransomware Recovery**	• OceanProtect®
Backup	• physical isolation within a protected part of the data center • Separate security credentials and multi factor authentication • Air gap • Immutable snapshots • Locked retention policies	• Immutable snapshots • DataLock (WORM) • Air gap • RBAC • Multi-factor authentication	• Immutable snapshots • WORM • RBAC • Multi-factor authentication
Recovery	• Automated restore and recovery • Performing damage assessments • Remediate and remove the offending malware	• Fast Global Search across all workloads and locations • Instant Mass Restore recovers VMs and files • Large file recovery optimization	
Detect / Analysis	• CyberSense® (ML) to find data corruption	• Intrusion detection by Helios ML • Detect file-level anomalies	

CyberSense® provides full content indexing and uses machine learning (ML) to analyze over 100 content-based statistics and detect signs of corruption due to ransomware. CyberSense® finds corruption with up to 99.5% confidence, helping you identify threats and diagnose attack vectors while protecting your business-critical content within the security of the vault.

Helios ML algorithms proactively assess and monitors infrastructure resources regularly. If your data change rate, including data ingest is out of the normal range—based on daily change rates on logical data, stored data after global deduplication, or historical data ingest—Helios sends a notification to your IT administrators. Besides monitoring backup data change rate to detect a potential ransomware attack, Cohesity® detects and alerts for file-level anomalies within unstructured files and object data. This includes analyzing the frequency of files accessed, number of files being modified, added or deleted by a specific user or an application to ensure, a ransomware attack is quickly detected.

FUTUREWEI Technologies

References

- Dell® EMC® Data Domain Operating System (Dell®technologies.com)
- h12927-Dell®EMC®-PowerProtect®-dd-ss.pdf (Dell®technologies.com)
- Dell® EMC® Data Domain SISL™ Scaling Architecture (Dell®technologies.com)
- isolated-recovery-solution-overview.pdf (Dell®technologies.com)
- Cohesity® SpanFS™ and SnapTree™
- Cohesity®_Data_Protection_White_Paper.pdf
- Simplify Long-Term Data and Application Retention and Archival (Cohesity®.com)
- Cohesity®-DataProtect®-Datasheet.pdf
- C6000 | Cohesity®
- Architecture Matters (Cohesity®.com)

Thank You.