EMC DATA DOMAIN SYSTEMS

Deduplication Storage for Backup and Archive

ESSENTIALS

Fast Inline Deduplication

- Up to 31 TB/hr performance
- CPU-centric scalability

Scalable Deduplication Storage

- Reduces backup and archive storage requirements by 10 to 30 times
- Provides up to 100 PBs of logical storage for long-term backup retention

Data Invulnerability Architecture

- Inline write/read verification, continuous fault detection and healing
- Dual disk parity RAID 6

Easy Integration

- Supports leading enterprise backup and archive applications for database, file, email, virtual environments and more

Backup and Archive Consolidation

- Simultaneously supports backup and archive data
- Meets governance policies and compliance regulations for archive data

Fast, Efficient and Scalable Disaster Recovery

- Reduce bandwidth requirements by up to 99 percent
- Replicate from up to 270 remote sites into a single system

Operational Simplicity

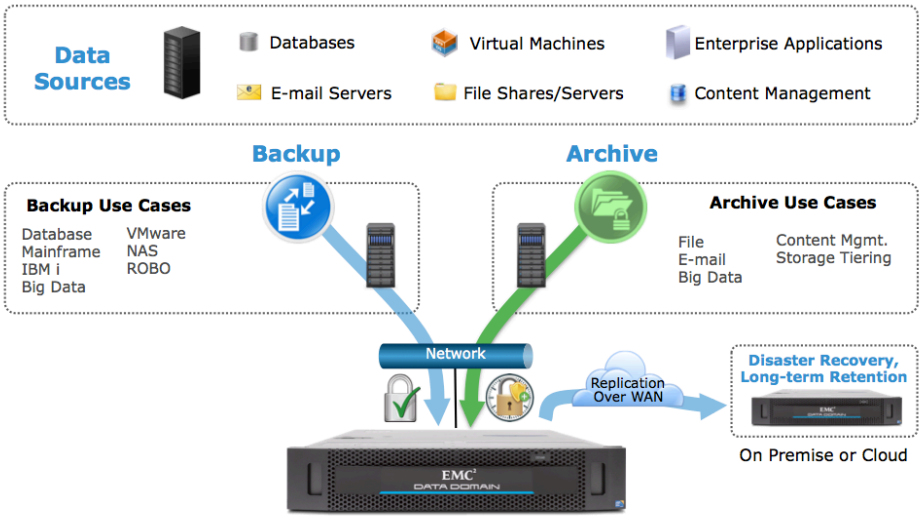
- Lower administrative costs
- Automatic call-home reporting
- Reduced footprint

Deduplication reduces the amount of disk storage needed to retain and protect data by an average of 10-30 times, making disk a cost-effective alternative to tape for backup and archive. Data on disk is available online and onsite for longer retention periods, and restores and retrievals become fast and reliable. Storing only unique data on disk also means that it can be cost-effectively replicated over existing networks to remote sites for safe, reliable tape-free disaster recovery.

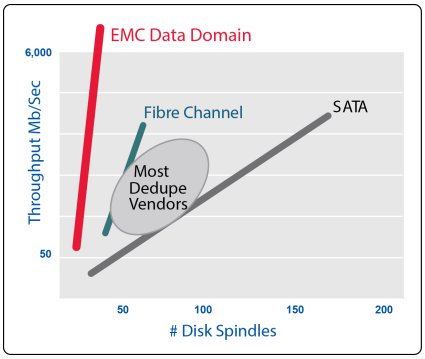
EMC Data Domain deduplication storage systems continue to revolutionize backup, archiving, and disaster recovery with high-speed, inline deduplication. Data Domain systems scale from smaller remote office appliances to highly scalable systems for backup and archive data at large enterprise data centers.

FAST INLINE DEDUPLICATION

As the industry's fastest deduplication storage, Data Domain systems provide throughput up to 31 TB/hour, enabling more backups to complete sooner, reducing pressure on backup windows. In order to minimize disk requirements, Data Domain systems deduplicate data inline—during the backup and archive process—so that the data lands on disk already deduplicated.



EMC Data Domain systems deduplicate data inline during the backup or archive process, which maximizes backup performance white minimizing disk storage requirements. Deduplicated data can be stored onsite for immediate restores and longer-term retention on disk. The deduplicated data can also be replicated over the WAN to a remote site for disaster recovery operations, eliminating the need for tape-based backups, or for consolidating tape backups to a central location. Data Domain systems provide the capability to consolidate both backup and archive data on the same infrastructure allowing greater consolidation by eliminating silos of storage and associated overhead.



CPU-Centric Storage
Data Domain Stream Informed Segment Layout (SISL) scaling architecture takes the pressure off of disk I/O as a bottleneck, so the remaining system design is CPU-Centric. Other deduplication methods require more disks to increase their throughput speeds.

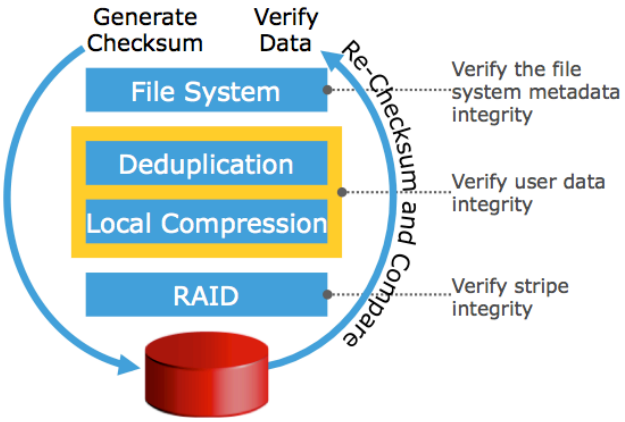
The key to enabling industry leading performance while minimizing disk requirements is the EMC Data Domain Stream-Informed Segment Layout (SISL™) scaling architecture. Specifically, SISL leverages the continued advancement of CPU performance to continuously increase Data Domain system performance by minimizing disk accesses required to deduplicate data. SISL deduplicates data inline by identifying duplicate data segments in memory, which minimizes disk usage. This enables Data Domain system throughput to be CPU-centric, not "spindle bound."

SCALABLE DEDUPLICATION STORAGE

By reducing backup and archive storage requirements by 10 to 30x, Data Domain systems can help significantly reduce the storage footprint for a typical enterprise data set. For environments that need to retain backups for long periods of time to meet legal mandates, the EMC Data Domain Extended Retention software option enables a logical system capacity of up to 100 PB for long-term retention of backup data. With deduplication, months of retention on disk is possible using the same number of "floor tiles" that traditionally provided only a couple of days of disk staging.

DATA INVULNERABILITY ARCHITECTURE

Data Domain systems are designed as the storage of last resort – built to ensure you can reliably recover your data with confidence. The EMC Data Domain Data Invulnerability Architecture is built into every Data Domain system to provide the industry's best defense against data integrity issues. Inline write and read verification protects against and automatically recovers from data integrity issues during data ingest and retrieval. Capturing and correcting I/O errors inline during the backup process eliminates the need to repeat backup jobs, ensuring backups complete on time and satisfy service-level agreements. In addition, unlike other enterprise arrays or file systems, continuous fault detection and self-healing ensures data remains recoverable throughout its lifecycle on a Data Domain system.



End-to-end data verification

End-to-end data verification reads data after it is written and compares it to what was sent to disk, proving that it is reachable through the file system to disk and that the data is not corrupted. Specifically, when the Data Domain Operating System receives a write request from backup software, it computes a checksum over the data. After analyzing the data for redundancy, it sotres the new data segments and all of the checksums. After all the data is written to disk, the Data Domain Operating System verifies that it can read the entire file from the disk platter and through the Data Domain file system, and that the checksums of the data read back match the checksums of the written data. This confirms the data is correct and recoverable from every level of the system.

Data Domain System Specifications

	DD160	DD620	DD640[3]	DD670[3]	DD860[3]	DD890[3]	DD990[3]
Logical Capacity[1,2]	40 - 195 TB	83 - 415 TB	.32 – 1.6 PB	0.6-2.7 PB	1.4 - 7.1 PB 5.7 - 28.5 PB[7]	2.9-14.2 PB	5.7 - 28.5 PB 20 - 100 PB[7]
Max. Throughput (Other)	667 GB/hr	1.1 TB/hr[4]	2.3 TB/hr[8]	3.6 TB/hr[5]	5.1 TB/hr[5]	8.1 TB/hr[6]	15.0 TB/hr[5]
Max. Throughput (DD Boost)	1.1 TB/hr	2.4 TB/hr	3.4 TB/hr	5.4 TB/hr	9.8 TB/hr	14.7 TB/hr	31.0 TB/hr

1. Mix of typical enterprise backup data (file systems, databases, email, developer files). The low end of capacity range represents a full backup weekly or monthly, incremental backup daily or weekly, to system capacity. The top end of the range represents full backup daily to system capacity. All capacity values are calculated using Base10 (i.e., 1TB = 1,000,000,000,000 bytes).
2. Requires EMC Data Domain Extended Retention software option.

EASY INTEGRATION

Data Domain systems integrate easily with existing infrastructures and can be used seamlessly with a variety of applications across both backup and archive workloads. Integrating a Data Domain system does not require any change in process or infrastructure, so you can realize the value of deduplication quickly and efficiently. All leading backup applications as well as the industry-leading archiving applications are supported with a Data Domain system. In addition, users can leverage a Data Domain system as the target for application protection utilities like Oracle RMAN or write directly over CIFS or NFS to support additional adjacent workloads.

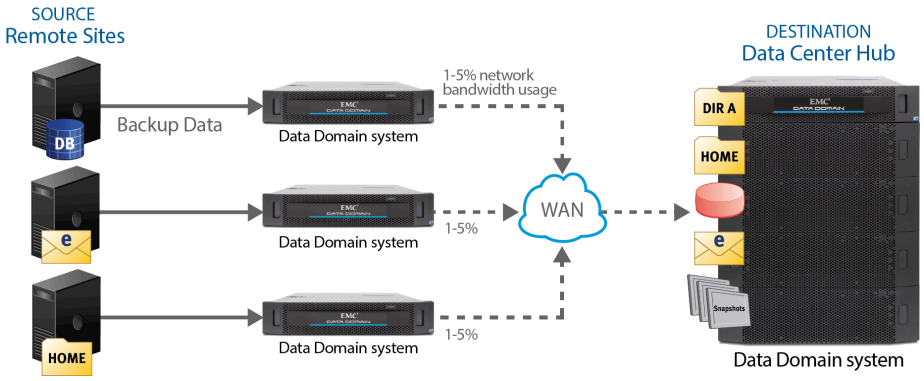
Since Data Domain systems simultaneously support multiple access methods including NFS, CIFS, VTL, NDMP and EMC Data Domain Boost™ all applications and utilities can be supported in the same Data Domain system, at the same time to enable greater storage consolidation. A system can present itself as a file server, offering NFS, CIFS access over Ethernet; as a virtual tape library (VTL) over Fibre Channel; as an NDMP tape server over Ethernet; or as a disk target using application specific interfaces like Data Domain Boost.

BACKUP AND ARCHIVE CONSOLIDATION

Data Domain systems are the first and only inline deduplication system to simultaneously support backup and archive data. This enables Data Domain systems to reduce overall total cost of ownership (TCO) by sharing resources across backup and archive data. Specifically, a single Data Domain system can be used for backup and recovery of the entire enterprise (including Oracle, Microsoft, and VMware®, as well as IBM i and mainframe environments) as well as archive workloads (including file, e-mail, and enterprise content management). Data Domain systems cost effectively protect archive data through integration with industry leading archiving applications including EMC SourceOne and Symantec Enterprise Vault. By consolidating to a common deduplication storage system, you can eliminate silos of storage and associated overhead - including management, floor space, power and cooling. In addition, with EMC Data Domain Retention Lock software, Data Domain systems can meet internal governance policies or compliance regulations for archive data including SEC 17a-4(f).

FAST, EFFICIENT AND SCALABLE DISASTER RECOVERY

As data lands on a Data Domain system, it will immediately begin replicating it to a disaster recovery site. To meet strict DR requirements, EMC Data Domain Replicator software can replicate at 52 TB/hr over a 10 Gb network connection. Data Domain systems only replicate unique compressed data across the network, requiring a fraction of the time, bandwidth and cost of traditional replication methods. With cross-site deduplication only unique data is transferred across any of the WAN segments. This can reduce WAN bandwidth requirements up to 99%, making network-based replication fast, reliable and cost-effective. For the highest level of security, data being replicated between Data Domain systems can be encrypted by using the standard Secure Socket Layer (SSL) protocol. Data Domain systems provide flexible replication topologies including full system mirroring, bi-directional, many-to-one, one-to-many, and cascaded. In a many-to-one deployment, data from up to 270 remote offices can be replicated to a single DD990 system.



CONTACT US

To learn more about how EMC products, services, and solutions can help solve your business and IT challenges, contact your local representative or authorized reseller—or visit us at www.EMC.com.

OPERATIONAL SIMPLICITY

Data Domain systems are very simple to install and manage resulting in lower administrative and operational costs. Administrators can access the Data Domain Operating System through command line over SSH or through EMC Data Domain System Manager, a browser-based graphical user interface. Initial configuration and updates can easily be made for multiple systems, along with the monitoring of system states and operations. Simple scriptability as well as SNMP monitoring provides additional management flexibility.

In addition, all Data Domain systems have automatic call-home system reporting called autosupport, which provides email notification of complete system status to EMC support and a selected list of administrators. This non-intrusive alerting and data collection capability enables proactive support and service without administrator intervention, further simplifying ongoing management.

Published in the USA. 4/13 Data Sheet H6811.7