

EMC DATA DOMAIN OPERATING SYSTEM

Powering EMC Protection Storage

The Data Domain Operating System (DD OS) is the intelligence that powers EMC Data Domain deduplication storage systems. It provides the agility, security and reliability that enables Data Domain systems to deliver scalable, high-speed protection storage for backup, archive and disaster recovery.

ESSENTIALS

HIGH-SPEED, SCALABLE DEDUPLICATION

- Up to 58.7 TB/hr performance
- Reduces protection storage requirements by 10 to 30x
- CPU-centric scalability

DATA INVULNERABILITY ARCHITECTURE

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

SEAMLESS INTEGRATION

- Supports leading enterprise backup and archive applications

BACKUP AND ARCHIVE CONSOLIDATION

- Efficiently protect backup and/or archive data
- Meets governance policies and compliance regulations for archive data

CLOUD-READY

- Secure multi-tenancy provides logical isolation of user data
- Comprehensive tenant-level management and reporting

FAST, EFFICIENT AND SCALABLE DISASTER RECOVERY

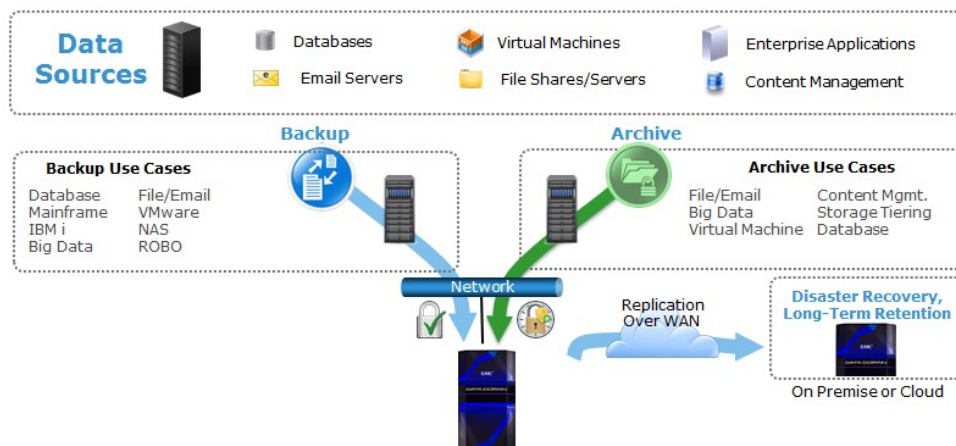
- Reduce bandwidth requirements by up to 99%
- Replicate from up to 540 remote sites into a single system

OPERATIONAL SIMPLICITY

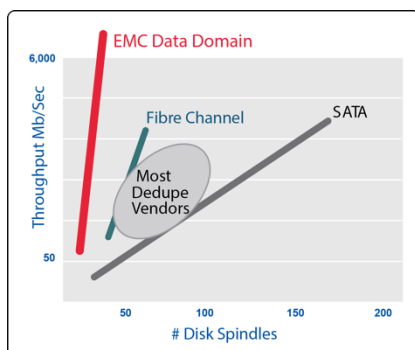
- Lower administrative costs
- Automatic call-home reporting
- Reduced footprint

HIGH-SPEED SCALABLE DEDUPLICATION

The Data Domain Operating System powers Data Domain systems to deliver industry-leading speed and efficiency with throughput up to 58.7 TB/hour, enabling more backups to complete sooner and reducing pressure on backup windows. DD OS employs variable-length deduplication to minimize disk requirements and ensure data lands on disk already deduplicated. This reduces backup and archive storage requirements by an average of 10 to 30x, making disk a cost-effective alternative to tape. Data on disk is available online and onsite for longer retention periods and restores and retrievals become fast and reliable. This efficiency enables Data Domain systems to protect up to 86.4 PB of data on a single system. 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.



EMC Data Domain Operating System powers EMC Data Domain deduplication storage systems to deduplicate data during either the backup process or archive process, which maximizes performance while 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 or a service provider site in the cloud 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 backup and archive storage and associated overhead.



CPU-Centric Storage

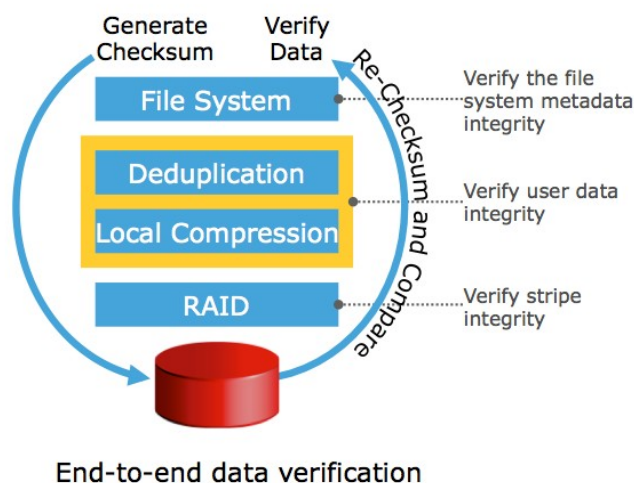
Data Domain Stream Informed Segment Layout (SISL) scaling architecture takes the pressure off of disk I/O as a bottleneck, so the remaining system design is CPU-Centric. Other deduplication methods require more disks to increase their throughput speeds.

The key that enables DD OS to provide 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 by identifying duplicate data segments in memory, which minimizes disk usage. This enables Data Domain system throughput to be CPU-centric, not “spindle bound.”

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 DD OS 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 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 stores 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 SYSTEMS

DATA DOMAIN CONTROLLER PERFORMANCE AND CAPACITY

	DD2200	DD2500	DD4200	DD4500	DD7200	DD9500
Maximum Throughput (Other)	3.8 TB/hr	5.6 TB/hr	10.6 TB/hr	10.6 TB/hr	12.6 TB/hr	27.7 TB/hr
Maximum Throughput (DD Boost)	4.7 TB/hr	13.4 TB/hr	25.6 TB/hr	25.6 TB/hr	28.3 TB/hr	58.7 TB/hr
Logical Capacity ¹	40-860 TB	1.3-6.6 PB	1.8-9.4 PB	2.8-14.2 PB	4.2-21.4 PB	8.6-43.2 PB
Logical Capacity w/ DD Extended			3.7 - 18.9 PB	5.7 - 28.5 PB	8.5 - 42.8 PB	17.2-86.4 PB

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).

SEAMLESS INTEGRATION

Data Domain systems integrate easily with existing infrastructures and can be used seamlessly with leading backup and archiving applications. Integrating a Data Domain system into your environment does not require any change in process or infrastructure, so you can realize the value of deduplication quickly and efficiently. In addition, Data Domain systems can integrate directly with leading enterprise applications such as Oracle RMAN or write directly over CIFS or NFS to support a variety of 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 protection 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.

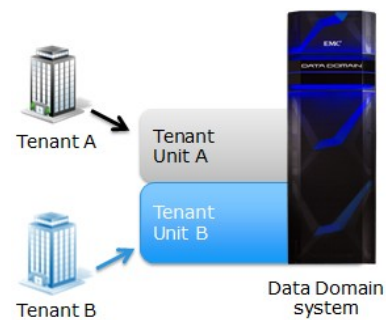
BACKUP AND ARCHIVE CONSOLIDATION

The flexibility of DD OS enables Data Domain systems to be the only protection storage 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, SAP, Microsoft, and VMware®, as well as IBM i and mainframe environments) as well as protecting archive data (including file, e-mail, enterprise content management, database and Virtual Machine archiving). 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 protection storage platform, 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).

CLOUD-READY

DD OS provides secure multi-tenancy, which enables large enterprises and service providers to deliver data protection as a service with Data Domain in a private or hybrid cloud. With secure multi-tenancy, a Data Domain system will logically isolate tenant data, ensuring that each tenant's data is only visible and accessible to them. Multi-tenant management and monitoring enables chargeback, trending, stream and capacity quotas and tenant-level reporting across multiple Data Domain systems. Secure multi-tenancy allows for tenant self-service by allowing tenants visibility only to the resources they are privileged to and catering to their needs for monitoring their resources.

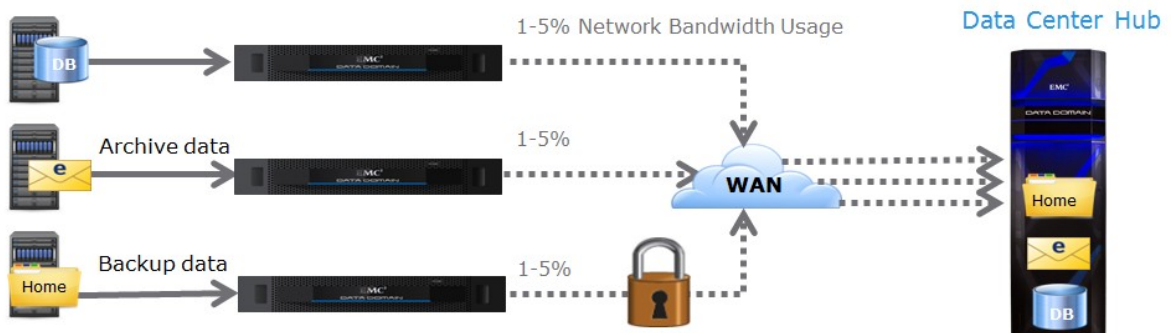
DD OS also offers physical capacity measurement to capture how much physical capacity is being consumed at a file, directory, MTree, tenant, or tenant-unit level. This serves as an effective mechanism for managing shared Data Domain protection storage capacity between individual departments or tenants.



FAST, EFFICIENT AND SCALABLE DISASTER RECOVERY

As data lands on a Data Domain system, it can 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. DD OS replicates only 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 540 remote offices can be replicated to a single DD9500 system.

Source Remote Sites



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.

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

VISIT THE EMC STORE

To explore Data Domain products.



store.emc.com/datadomain

www.EMC.com

EMC², EMC, the EMC logo, are registered trademarks or trademarks of EMC Corporation in the United States and other countries. VMware are registered trademarks or trademarks of VMware, Inc., in the United States and other jurisdictions. All other trademarks used herein are the property of their respective owners. © Copyright 2015 EMC Corporation. All rights reserved. Published in the USA. 10/15 Data Sheet H6811.9

EMC believes the information in this document is accurate as of its publication date. The information is subject to change without notice.