

EMC DATA DOMAIN BOOST SOFTWARE

Increase deduplication storage performance with advanced application integration

ESSENTIALS

Significant Reduction in Backup Time

- Up to 31 TB/hr aggregate throughput performance
- Distributed deduplication process dramatically increases throughput
- Reduced network bandwidth utilization
- Efficient CPU utilization on backup servers
- Faster restarts of failed backup jobs
- Enables faster “time-to-DR” readiness

Seamless Integration

- Centralized management and tighter integration with applications
- Heterogeneous backup server environment support

Replication Managed from Application Console

- 99 percent bandwidth reduction
- Cost-efficient disaster recovery
- Multisite tape and disk consolidation
- Encrypted replication

Advanced Load Balancing and Link Failover

- Scalable link aggregation at the application layer
- Simplified backup application configuration
- Seamless load balancing of jobs among available ports
- Link failover keeps backups operational

Ultra-Safe Storage for Reliable Recovery

- Continuous recovery verification, fault detection, and healing
- End-to-end data integrity

NEXT-GENERATION BACKUP AND RECOVERY

Backup applications are a critical component of backup, recovery, and disaster preparedness strategies and often direct the operation of disk and tape media that typically store five to ten times more data than can be found on the primary storage in the same data center. As data continues to grow unabatedly and as more attention is focused on efficient backup and disaster recovery (DR), traditional approaches to backup are no longer acceptable. Disk-based deduplication solutions are becoming mandatory to meet the backup and recovery challenges in today’s data center.

These transformational backup and DR solutions must have a foundation of simplicity and flexibility that enable consolidated management from a single pane within a backup application. Unfortunately, many disk-based backup technologies have only focused on improving the speed and reliability of backups, not advancing DR and ease-of-management. For example, some virtual tape libraries (VTLs) actually worsened the complexities and liabilities of tape-based backup and recovery. With many traditional VTLs, users have to provision virtual tape drives, slots, and cartridges, which could increase management complexity. The inability of data management tools to handle multiple instances of the same tape barcode or image makes management of DR copies a breaking point.

EMC® Data Domain® deduplication storage systems offer an alternative to traditional VTLs with a simple NFS/CIFS interface. Data Domain systems are simple to integrate with existing backup software, and allow users to enjoy the retention and recovery benefits of inline deduplication as well as the offsite disaster recovery protection of replication over the wide area network (WAN).

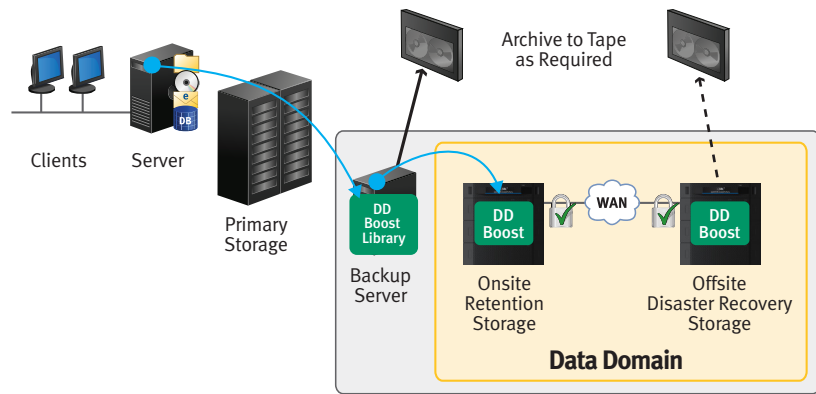
EMC Data Domain Boost extends the optimization capabilities of Data Domain systems for environments with EMC Avamar®, EMC NetWorker®, EMC Greenplum®, Oracle RMAN, Quest vRanger, and Symantec OpenStorage. DD Boost significantly increases performance by distributing parts of the deduplication process to the backup server or application clients, simplifies disaster recovery procedures, and serves as a solid foundation for additional integration between backup applications and Data Domain systems.

SIGNIFICANT REDUCTION IN BACKUP TIME

Without DD Boost, the backup server or application client will send all data, unique or redundant, to a Data Domain system for deduplication processing. With DD Boost, parts of the deduplication process are distributed to the backup server or application clients, enabling it to send only unique data segments to a Data Domain system. This dramatically increases the aggregate throughput, up to 31 TB/hour, and reduces the amount of data transferred over the network by 80 to 99 percent. These efficiencies can help eliminate future costs by leveraging existing backup servers and Ethernet networks.

EMC Data Domain Boost

Data Domain Boost significantly increases performance by distributing parts of the deduplication process to the backup server or application clients, simplifies disaster recovery procedures, and serves as a solid foundation for additional integration between backup applications and Data Domain systems.



In addition to performance improvements and network bandwidth benefits, the reduction in data transferred also decreases CPU utilization on the backup servers or application clients since sending data over the network is significantly more CPU intensive than the distributed deduplication process. Consequently, this improved efficiency provides an ability to drive more backups with existing resources or simplify the environment by removing a backup server to backup directly from the client.

DD Boost also increases the speed of restart and completion of failed backups. Since only unique data is sent over the network, once a failed job restarts, the data that has already been sent to the Data Domain system for a given backup job does not need to be sent again. This not only reduces the load on the network substantially, but also improves the overall throughput for the failed backups upon retry.

Overall, DD Boost increases aggregate throughput, substantially reduces backup windows, and improves backup server and application client efficiency. Best of all, DD Boost is transparent to the backup application because the distributed segment processing is handled by DD Boost on the backup server or application client.

SEAMLESS INTEGRATION

The combination of a Data Domain system and an application that supports DD Boost creates an optimized connection to provide a tightly integrated solution. DD Boost offers operational simplicity by enabling the backup application to manage the connection between the backup server or application client and one or more Data Domain systems. Unlike a VTL implementation, DD Boost does not require any of the artifacts of tape or tape emulation.

Similar to the Data Domain NFS/CIFS implementation, backup images can be simultaneously written to and read from the Data Domain system in the applications' native image formats. Backup images can be expired one at a time, leading to improved storage efficiency over VTL-only implementations.

DD Boost simplifies the management and allows easy sharing and simultaneous use of disk pools among multiple heterogeneous backup servers or application clients. This facilitates load balancing and the use of the "best" available backup servers for completing a particular backup job based on pre-configured policies.

DD Boost works on multiple backup server and client operating systems providing a seamless experience to administrators in heterogeneous environments. Further, DD Boost serves as a consolidated interface to all Data Domain systems making the systems transparent to the backup application. This simplifies the management of Data Domain deduplication storage systems.

MANAGE FILE REPLICATION FROM THE APPLICATION CONSOLE

Network-efficient replication technology sends backup images from one Data Domain system to one or more systems, when all systems are configured with DD Boost software. Using this approach, the bandwidth required for replication is reduced by up to 99 percent. This dramatically reduces the time and WAN bandwidth needed to create multiple copies of backups for tape consolidation or disaster recovery purposes. Resource usage on the backup servers or application clients is also reduced as they are no longer in the data path when duplicate copies of the backup are being created and transferred.

With DD Boost, applications can control replication between multiple Data Domain systems and provide backup administrators with a single point of management for tracking all backups and duplicate copies. This paradigm allows backup administrators to efficiently create DR copies of their backups over the WAN using EMC Data Domain Replicator software, and keep track of all the copies in the backup application's catalog for easy disaster recovery. This also enables administrators to manage retention periods for various copies individually, resulting in a more flexible DR deployment. If confidentiality is required, deduplicated and compressed data can be encrypted in-flight when being replicated between Data Domain systems, independent of the replication topology used.

Leveraging WAN-efficient Data Domain replication, DD Boost eliminates the need for tape-based backup and recovery procedures at remote sites. Recovery from backup copies at the central site also becomes simplified because all copies are tracked in the application's catalog. Since tape is no longer required at the remote sites, one can deploy Data Domain systems with DD Boost software as the foundation of a multi-site tape consolidation strategy.

ADVANCED LOAD BALANCING AND LINK FAILOVER

The advanced load balancing and link failover feature allows for aggregating multiple Ethernet links on the Data Domain system into a single group. This makes multiple links appear as one to the backup application. The system then transparently balances the backup load between Ethernet links in the group. DD Boost now supports multiple groups on a single Data Domain system to provide application level load balancing and failover in environments with multiple subnets. Link aggregation simplifies management of multiple backup policies and provides full utilization of all available Ethernet links. Compared to link aggregation at the Ethernet level, this creates a more seamless and scalable solution with higher network bandwidth utilization.

In addition, failover capabilities assist when one of the interfaces in the group goes down while the Data Domain system is still operational. The Data Domain system will automatically route both in-flight and subsequent backup jobs from the failed interfaces to the surviving interfaces. This automatic link failover mechanism keeps the backup systems operational in case of temporary network glitches.

Advanced load balancing and link failover works seamlessly with the underlying network connectivity and supports both physical and virtual interfaces. Further, unlike aggregation at the Ethernet layer, there is no special configuration required on the Ethernet switches.

ULTRA-SAFE STORAGE FOR RELIABLE RECOVERY

Systems with DD Boost benefit from the EMC Data Domain Data Invulnerability Architecture, which provides 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. Unlike other enterprise arrays or file systems, continuous

fault detection and self-healing features protect data throughout its lifecycle on all Data Domain systems.

DD Boost extends this protection to the backup server or application client by generating checksums on the data that it sends to the Data Domain system. The checksums are then transferred along with the data. The Data Domain system receiving the data computes new checksums on the incoming data and then compares them to the computed values from DD Boost for verification purposes. This ensures end-to-end verification of data.

Feature	Benefit
Distributed segment processing	Significant reduction in backup time Up to 31.0 TB/hr aggregate throughput performance Improved network bandwidth utilization Efficient CPU utilization on backup servers and application clients Less time required to restart failed backup jobs
Seamless integration	Centralized management and operational simplicity Backup server load balancing Heterogeneous backup server environment support Facilitates hierarchical duplication and centralized tape out operations
Backup application managed replication	Network efficient replication using DD Replicator 99 percent bandwidth reduction Cost-efficient disaster recovery Multi-site tape consolidation Encryption of replication session between Data Domain systems
Advanced load balancing and link failover	Scalable link aggregation at the application layer Simplified configuration Optimize throughput of multiple links Link failover keeps backups operational in case of temporary network glitches and failures
Data Invulnerability Architecture	Ultra-safe storage for reliable recovery End-to-end data integrity

SPECIFICATIONS

SOFTWARE

EMC Data Domain Operating System 5.2 or later
EMC Data Domain Boost Library 2.4 or later
EMC Data Domain Extended Retention software
EMC Data Domain Replicator software
EMC Avamar 6.0 or later
EMC NetWorker 7.6.1 or later
EMC Greenplum Database 4.2 or later
Oracle Recovery Manager
Quest vRanger 5.4 or later
Symantec NetBackup 6.5 or later
Symantec Backup Exec 2010 or later

HARDWARE

EMC Data Domain Appliance Series

BACKUP SERVER PLATFORMS

Oracle Enterprise Linux 5/6 (x86_64)
Solaris 9/10 (SPARC)
Solaris 10 (x86_64)
Red Hat Enterprise Linux 4/5 (x86, x86_64)
SuSE Linux Enterprise Server 10/11 (x86, x86_64)
Windows Server 2003 (x86, x86_64)
Windows Server 2008 (x86_64)
Windows Server 2008 R2 (x86_64)
HP-UX 11.23/11.31 (PA-RISC)
HP-UX 11.23/11.31 (IA64)
AIX 5.3/6.1 (POWER)

NETWORK CONNECTIVITY

IP connectivity between backup servers and Data Domain systems

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.

EMC², EMC, Avamar, Data Domain, Greenplum, NetWorker, and the EMC logo are registered trademarks or trademarks of EMC Corporation in the United States and other countries. All other trademarks used herein are the property of their respective owners. © Copyright 2011, 2012 EMC Corporation. All rights reserved. Published in the USA. Data Sheet 05/12 H7034.5

EMC Corporation
Hopkinton, Massachusetts 01748-9103
1-508-435-1000
In North America 1-866-464-7381
www.EMC.com

EMC Backup Recovery Systems
Santa Clara, California 95054
1-408-980-4800
In North America 1-866-933-3873

EMC²