

SLOW BACKUPS GOT YOU DOWN?

The Business Value of Data Domain Boost

WHY YOU SHOULD TAKE THE TIME TO READ THIS PAPER

This paper explains the technical differentiation of Dell EMC Data Domain Boost and why it matters to customers.

- **Increase backup speeds by up to 50%** (Finish backups within backup windows with breathing room for data growth. With performance up to 68 TB/hr, we are about 1.5 times faster than our nearest competitor.)
- **Use up to 99% less bandwidth** (Use your existing network infrastructure more efficiently.)
- **Improve disaster recovery** (Replace tape-based Disaster Recovery (DR) with bandwidth efficient replication improving performance, reliability, and time-to-DR readiness with simplified DR testing.)
- **Increase backup success** (Maximize success through link aggregation & failover with fewer failed backup jobs to restart.)
- **Simplify day-to-day operations** (Eliminate the headaches and costs of managing thousands of physical or virtual tape cartridges.)
- **Reduce backup server workload by up to 40%** (Reduce the CPU workload on backup servers or application clients.)
- **Improve catalog visibility & control** (Achieve full backup catalog awareness for DR.)

October, 2016

The information in this publication is provided “as is.” Dell EMC makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any Dell EMC software described in this publication requires an applicable software license.

EMC², EMC, the EMC logo, are registered trademarks or trademarks of Dell EMC in the United States and other countries. All other trademarks used herein are the property of their respective owners. © Copyright 2016 Dell EMC. All rights reserved. Published in the USA., 10/16; white paper, H11755.1

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

EMC is now part of the Dell group of companies.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	5
Have slow backups got you down?	5
Introduction	5
Audience	5
PROBLEMS WITH TRADITIONAL BACKUPS.....	6
Factors that limit backup performance and success.....	6
There must be a better way	6
DATA DOMAIN SISL – FOUNDATION FOR SPEED.....	6
WHAT IS DATA DOMAIN BOOST?.....	7
DD Boost uses an optimized transport.....	7
DD Boost distributes the deduplication process	7
DD Boost with managed file replication.....	8
DD Boost provides advanced load balancing and failover.....	8
THE BENEFITS OF DD BOOST	8
Speed up backups by up to 50%.....	8
Maximize performance through path load balancing	8
Maximize backup success with automatic path failover.....	8
Use up to 99% less bandwidth	8
Reduce backup server workload by up to 40%	9
Improve disaster recovery.....	9
Simplify operations by eliminating tape cartridge management.....	9
Provide backup administrators full catalog visibility and control	9
Give enterprise application owners control of their own backups and restores.....	10
Minimize full backups with virtual synthetic backups	10
DD BOOST SOLUTION INTEGRATION.....	10
Data Domain Boost for Backup Applications.....	10
Data Domain Boost for Dell EMC Avamar	10
Data Domain Boost for Dell EMC Networker	11
Data Domain Boost for Veritas NetBackup and Backup Exec.....	12
Data Domain Boost for vSphere Data Protection Advanced	12
Data Domain Boost for HP Data Protector.....	12
Data Domain Boost for Veeam.....	12

Data Domain Boost for CommVault Simpana.....	12
Data Domain Boost for Pivotal Greenplum	12
Data Domain Boost for Enterprise Applications.....	12
Data Domain Boost for Oracle RMAN.....	13
Data Domain Boost for Microsoft SQL	14
Data Domain Boost for SAP	14
Data Domain Boost for SAP HANA.....	14
Data Domain Boost for IBM DB2.....	14
Data Domain Boost for Hortonworks and Cloudera Hadoop.....	14
Data Domain Boost Ecosystem.....	16
DD Boost file system plug-in (BoostFS)	16
CONCLUSION	17

EXECUTIVE SUMMARY

HAVE SLOW BACKUPS GOT YOU DOWN?

Traditional backups are just too slow, inefficient, and unreliable for the typical data center of today. The data being protected is growing dramatically and backup windows are starting to shrink. How can you effectively manage your backups when IT budgets are flat or declining? If it's bad now, it's only going to get worse with more servers and more data to protect. Someone once said, the definition of insanity is continuing to do the same thing over and over and expecting different results. It's time to stop the insanity with your backup strategy.

What you need is a faster, more efficient way to do your backups. What you need is a Dell EMC Data Domain deduplication storage system with Dell EMC Data Domain Boost™. DD Boost is an innovative technology that dramatically increases backup performance, improves backup reliability, improves operational and disaster recovery, and allows you to better leverage your current infrastructure investments. EMC was first to introduce distributed deduplication technology with DD Boost in 2010. Now, widespread integration of DD Boost over Ethernet or Fibre Channel for both LAN and WAN implementations means there are many ways to leverage this unique technology to improve your backup and recovery operations.

INTRODUCTION

This paper focuses on Data Domain Boost and the many benefits it provides compared to doing traditional backups. It is not intended to explain in detail how it works. The purpose of this paper is to explain the technical and business reasons why DD Boost will change the way you do backups.

AUDIENCE

This paper is intended for Dell EMC customers, Dell EMC sales, Dell EMC systems engineers, Dell EMC partners, and anyone else who is interested in learning more about Data Domain Boost technology and all the unique benefits that it can provide.

PROBLEMS WITH TRADITIONAL BACKUPS

FACTORS THAT LIMIT BACKUP PERFORMANCE AND SUCCESS

There are many factors that limit overall backup performance, which can cause backups to run into production windows, reduce backup success, or increase infrastructure costs. We will focus on the most common items:

- Limited LAN/SAN Bandwidth
- Inefficiency of CIFS and NFS for backups
- Limited number of Fibre Channel connections
- Speed of Fibre Channel connections
- Overworked backup servers
- Limited number of physical tape drives
- Unreliability of physical tape drives
- Tape media failures
- Slow tape drives

Many backups are limited to some degree by LAN/SAN bandwidth available to the server being backed up. Now add the backup overhead inefficiency of the CIFS and NFS protocols over Ethernet, and you have exacerbated the problem further. Other backups are limited by the speed and number of available Fibre Channel connections due to the expense of having to upgrade to the latest port speeds every few years. Another common choke point for traditional backups is overworked backup servers, especially during full backup periods.

Frequently, with traditional backups, you have to deal with a limited number of slow and unreliable tape devices that restrict backup performance. In addition, tape media failures reduce backup success and cause additional work for your staff to rerun failed backup jobs. And finally, with physical tape drives, if you can't keep the data streaming fast enough, the tape buffer empties and the tape drive dramatically slows down. This reduces overall backup performance and also causes premature wearing of the tape drive head, further reducing backup and restore reliability. Backups using physical tape or virtual tape also create the additional challenge of managing thousands of tape cartridges associated with day to day backups and disaster recovery.

There is a lot of change occurring in the data protection landscape today. Even with Purpose Built Backup Appliances or PPBA's, simply replacing tape is not enough. For application owners such as DBAs, difficulty meeting backup windows is becoming a problem. Due to an explosion of data that needs to be protected, and lack of control and visibility into their own backups, DBAs are becoming dissatisfied with their current solution. However, now the application owners have other alternatives for protecting their data. So when they don't get the control and visibility they want over their own backups, they go off on their own - leaving the IT team with less visibility and control than ever before. As a result, this can lead to unnecessary silos of storage, or what can be called, an accidental architecture.

THERE MUST BE A BETTER WAY

You may be thinking - I wish I could get better backup performance, but CIFS and NFS just aren't fast enough and my current Fibre Channel infrastructure is too old and slow and 10GB Ethernet has become the standard. Isn't there a way to dramatically improve backups using my existing infrastructure? Yes! That is exactly what Dell EMC has developed. It's called Data Domain Boost and it does some amazing things to improve your backups and disaster recovery!

DATA DOMAIN SISL – FOUNDATION FOR SPEED

Before discussing DD Boost, let's first touch on the original Data Domain innovation designed for speed. The foundation for Data Domain system's industry leading performance is its Stream Informed Segment Layout (SISLTM) scaling architecture. SISL enables Data Domain systems to perform 99% of the deduplication effort in CPU and RAM. This means that Data Domain systems are CPU-centric not spindle-bound for performance like other deduplication platforms. This is also the reason why Data Domain systems have dramatic increases in overall performance with each successive generation using the latest Intel processors.

WHAT IS DATA DOMAIN BOOST?

Dell EMC Data Domain Boost is a software option available for all Data Domain systems. DD Boost is made up of two components – a DD Boost plug-in that runs on the backup server or client and a DD Boost component that runs on the Data Domain system. All connectivity between components uses industry standard Ethernet or Fibre Channel. DD Boost software enables tight integration with backup and enterprise applications using an optimized transport. DD Boost includes three main features.

- **Distributed segment processing**, which distributes parts of the deduplication process from the Data Domain system to the backup server or client, increasing backup application performance by up to 50%.
- **Managed file replication**, which allows backup applications to manage Data Domain replication with full catalog awareness.
- **Advanced load balancing and link failover**, which provides link aggregation for load balancing and link failover, which eliminates the need for network layer aggregation.

With DD Boost and managed file replication, backup administrators can control the backup, restore, and replication of the backup images written to Data Domain systems from the backup application console in a catalog-aware manner. This ability gives administrators a single pane view for all the backups and copies of those backups.

DD Boost distributes the deduplication process to the backup server or client, simplifies disaster recovery procedures by providing WAN efficient Data Domain Replicator control to backup applications for creating multiple copies, and serves as a solid foundation for additional integration between backup applications and Data Domain systems.

In the next section, let's examine each of the DD Boost components in a little bit more detail.

DD BOOST USES AN OPTIMIZED TRANSPORT

There are components of the CIFS and NFS protocols that we really don't need for backups over Ethernet to a Data Domain system. These components that we don't need slow down the transport of data and can limit backup performance. DD Boost was designed to leverage only what we need to reliably move data over Ethernet and FC, and eliminate all the other overhead. Therefore, backups with DD Boost are faster and more efficient than with CIFS, NFS, or VTL.

DD BOOST DISTRIBUTES THE DEDUPLICATION PROCESS

Data Domain Boost 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 with managed file replication also gives the backup application control over replication. The more backup servers or clients you have, the more impactful DD Boost can be. For example, with DD Boost an environment with five or more backup servers would have five backup servers and the Data Domain system each doing some of the deduplication effort. Without DD Boost, the entire deduplication effort is being performed entirely by the Data Domain system and all the data must travel from the client to the Data Domain system. With some backup applications, the deduplication can be distributed all the way to the application or database servers.

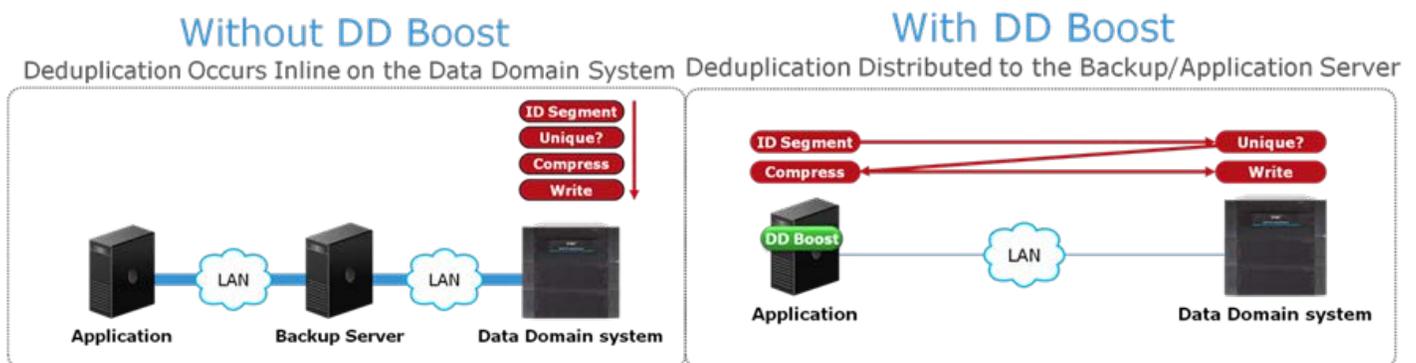


Figure 1: DD Boost distributed segment processing

DD BOOST WITH MANAGED FILE REPLICATION

DD Boost with managed file replication enables the backup application to manage replication of backup images from one Data Domain system to one or more other Data Domain systems with total catalog awareness of all local and remote copies.

DD BOOST PROVIDES ADVANCED LOAD BALANCING AND FAILOVER

DD Boost also provides advanced load balancing and failover across all available data paths without the need for network layer aggregation. This feature provides higher performance through more efficient utilization of existing network bandwidth and more reliable backups through path failover without the need to restart failed backup jobs due to a path failure.

THE BENEFITS OF DD BOOST

Sure, DD Boost seems impressive so far, but what does it do for my backup operations? Aren't all deduplication solutions pretty much the same? Don't they all pretty much offer the same technology advantages? NO! Data Domain systems with SISL and DD Boost provide tremendous advantages. They will save you time and money by using your current backup infrastructure more efficiently, simplify your day-to-day operations, and dramatically reduce your backup times.

SPEED UP BACKUPS BY UP TO 50%

By distributing the deduplication process, DD Boost dramatically speeds backups by up to 50% using your existing backup infrastructure. This is a simple, but very important point. Would you like to perform your backups in 8 hours without DD Boost, or in a little over 5 hours using the same exact hardware with DD Boost? In fact, with DD Boost, a DD9800 provides about 1.5 times faster performance than our nearest competition achieving backup speeds up to 68 TB/hr!

MAXIMIZE PERFORMANCE THROUGH PATH LOAD BALANCING

Data Domain system implementations typically leverage multiple paths from the backup servers to the Data Domain system. One of the additional benefits of DD Boost is automatic load balancing of the backup workload across all available paths, further maximizing backup performance and efficiency. This means you will be getting the full benefit of your existing network infrastructure through DD Boost without having to setup any special external load balancing networking options.

MAXIMIZE BACKUP SUCCESS WITH AUTOMATIC PATH FAILOVER

Another benefit of DD Boost with multiple backup servers to a Data Domain system is automatic path failover in the case of one path having a problem improving the reliability of your backups. This also means your backups continue to run if you lose a backup path, keeping your overall backup completion success percentage higher and fewer failed backup jobs that need to be restarted by your operations staff.

USE UP TO 99% LESS BANDWIDTH

Without DD Boost, when you are performing full backups, all data must be sent from the backup server to the Data Domain system. With DD Boost, only unique data is sent from the backup server or client to the Data Domain system. This means up to 99% less data has to be moved across the network – even for full backups. This allows you to more efficiently use your existing LAN or SAN resources. When DD Boost can be leveraged at the client (with NetWorker, Avamar, HP Data Protector, vSphere VDP, Veeam, Oracle RMAN, Microsoft SQL, IBM DB2, SAP, SAP HANA, Pivotal Greenplum, NetVault), this bandwidth advantage spans the entire backup path all the way from the client to the Data Domain system. For backup environments that may be experiencing bandwidth choking during full backup times, this can provide serious performance improvements and may defer the costs of doing infrastructure upgrades.

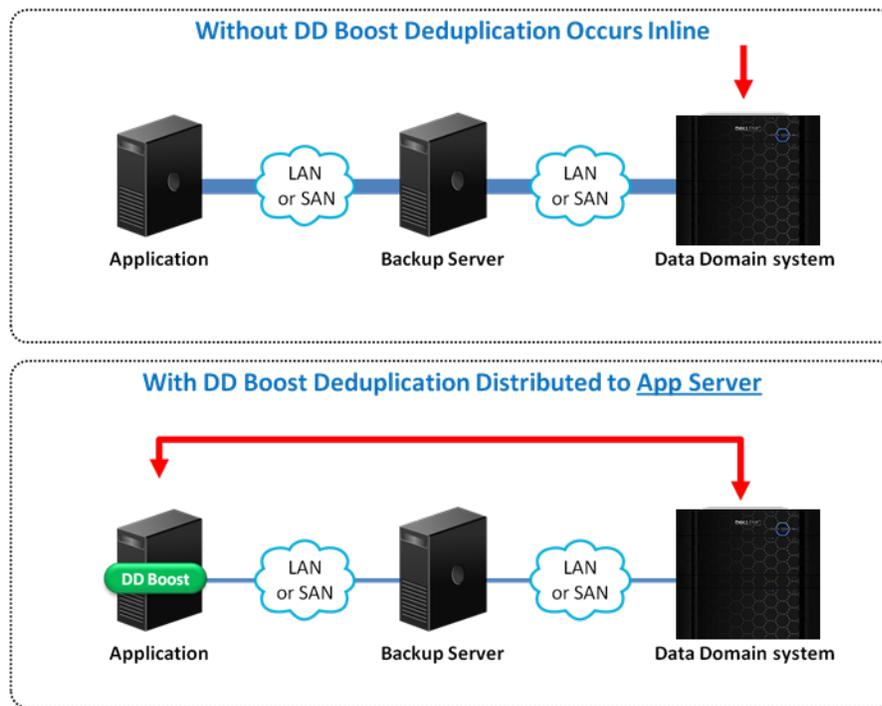


Figure 2: DD Boost minimizes bandwidth used

REDUCE BACKUP SERVER WORKLOAD BY UP TO 40%

You may be thinking that moving some of the deduplication heavy lifting from the Data Domain system to the backup servers or client is going to negatively impact your server workload. Good news - that is not the case! This may seem counterintuitive, but as it turns out, sending data is a very resource intensive process for a server. Therefore, sending less data significantly reduces the load on that server. In other words, it takes fewer CPU cycles to assist with 2 steps of the deduplication process than it takes to push full backups over Ethernet. A-ha, it's starting to sound very interesting isn't it? With DD Boost, your backups run faster, you use less bandwidth and we have reduced the workload for your backup servers. Wow! But wait, there's more.

IMPROVE DISASTER RECOVERY

DD Boost with Data Domain Replicator for disaster recovery provides the opportunity to replace physical tape and all the associated headaches and risks. There are no physical tapes to recall and wait for. There are no physical tapes to get fatal parity errors. There are no physical tapes to get damaged. There are no physical tapes that can get lost. You won't destroy backup tapes with faulty tape drives. All of this means you will have a more reliable and cost effective infrastructure for disaster recovery. Your recovery will also not be limited by a small number of physical tape drives at the DR site or the wasted time of loading, mounting and positioning each data cartridge before data can actually be recovered.

Data Domain Replicator ensures network-efficient replication of only unique data to one or more target sites providing the fastest time-to-DR readiness. This means that your replication bandwidth costs will be minimal and your time to data access for disaster recovery is fast and reliable.

If you are using Data Domain with CIFS, NFS, or DD Boost, you won't have any virtual tape cartridges to worry about. Furthermore, with DD Boost and managed file replication, your backup catalog will already be fully aware of all replicated copies available at your DR site.

SIMPLIFY OPERATIONS BY ELIMINATING TAPE CARTRIDGE MANAGEMENT

With DD Boost over Ethernet or Fibre Channel you will finally be able to eliminate the need to manage thousands of physical or virtual tape cartridges greatly simplifying your day-to-day production and disaster recovery operations. This dramatically reduces or eliminates the time, effort, and costs associated with handling and managing tape cartridges.

PROVIDE BACKUP ADMINISTRATORS FULL CATALOG VISIBILITY AND CONTROL

Backup administrators want complete control and visibility into their backups. Without DD Boost, the backup application is only aware of the local backup data in the Data Domain system and is unaware of any remote copies that might have been made by Data Domain Replicator. Additionally, without DD Boost, replicated data will have the same retention policies as the original data. DD Boost allows different retention to be set on replicated copies than the original source data. DD Boost with managed file replication provides visibility into backups being done and control over what is replicated from a local Data Domain system to 1 or more remote Data Domain systems. DD Boost with managed file replication allows backup applications to have total backup catalog awareness of all local and replicated copies at other sites.

GIVE ENTERPRISE APPLICATION OWNERS CONTROL OF THEIR OWN BACKUPS AND RESTORES

Many enterprise application owners and database administrators want to manage their own backups and restores and don't want to depend on backup administrators to do everything for them. DD Boost for Enterprise Applications gives enterprise application owners the ability to perform their own backups & restores and benefit from the DD Boost performance, reliability, and bandwidth efficiency advantages mentioned previously. Please refer to the integration section for more information.

MINIMIZE FULL BACKUPS WITH VIRTUAL SYNTHETIC BACKUPS

DD Boost supports virtual synthetic file system backups with Dell EMC NetWorker and Veritas NetBackup on Data Domain systems. Virtual synthetic backups reduce the processing overhead associated with traditional synthetic full backups. The Data Domain system uses metadata to synthesize a full backup without moving data over the network. Virtual synthetics enables the Data Domain system to synthesize a full backup from a previous full backup plus subsequent incremental backups, removing this burden from the backup server. This dramatically reduces data movement and eliminates the need to perform regular full backups of file systems every time.

- Backup admin can take fulls and incrementals and direct Data Domain to perform a virtual full

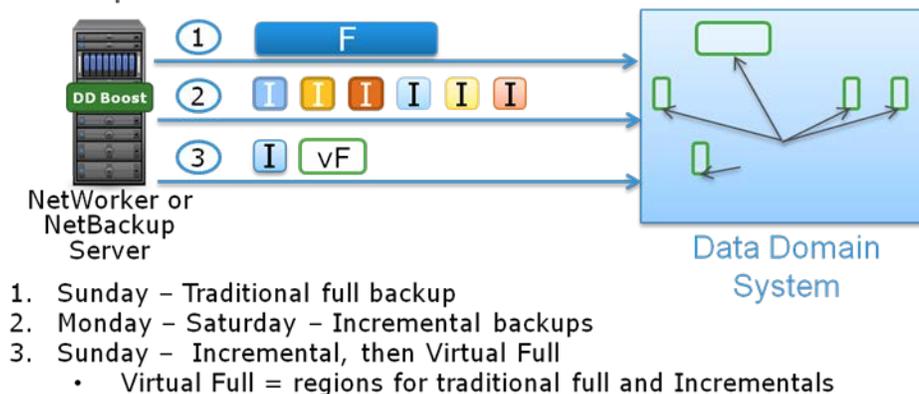


Figure 3: Virtual Synthetic Backups

DD BOOST SOLUTION INTEGRATION

Dell EMC continues to invest in the DD Boost ecosystem to ensure that as many customers as possible can experience the benefits of DD Boost! We have the broadest and deepest integration of its kind in the industry! This next section provides a high level explanation of our current DD Boost integration.

DATA DOMAIN BOOST FOR BACKUP APPLICATIONS

DD Boost has been integrated with many backup applications including Dell EMC Avamar, Dell EMC NetWorker, Veritas NetBackup, Veritas BackupExec, vRanger, NetVault, Commvault Simpana, HP Data Protector, Veeam, and VMware vSphere Data Protection Advanced providing all the previously mentioned DD Boost benefits. Each application leverages DD Boost to improve Data Domain backup performance, reliability and bandwidth efficiency. The next section covers each backup application's specific integration with DD Boost in more detail.

DATA DOMAIN BOOST FOR DELL EMC AVAMAR

DD Boost for Avamar enables the Avamar client to send backup data directly to the Data Domain system. Specifically, this integration provides the Data Domain system scalability and performance advantages for all backup workloads including remote offices over WAN, desktops and laptops, VMware image backups, NDMP, file systems, and enterprise applications like Oracle databases, Microsoft Exchange, Microsoft SharePoint, and Microsoft SQL databases. This means the up to 99% network bandwidth reduction advantage of DD Boost extends all the way from the client to the Data Domain system. It also means that you can now multiply the advantage of distributing some of the deduplication effort to hundreds of clients improving overall backup performance even more. DD Boost for Avamar also supports Avamar Instant Access to virtual machines stored on Data Domain.

Avamar with Data Domain Systems

All Workloads Supported

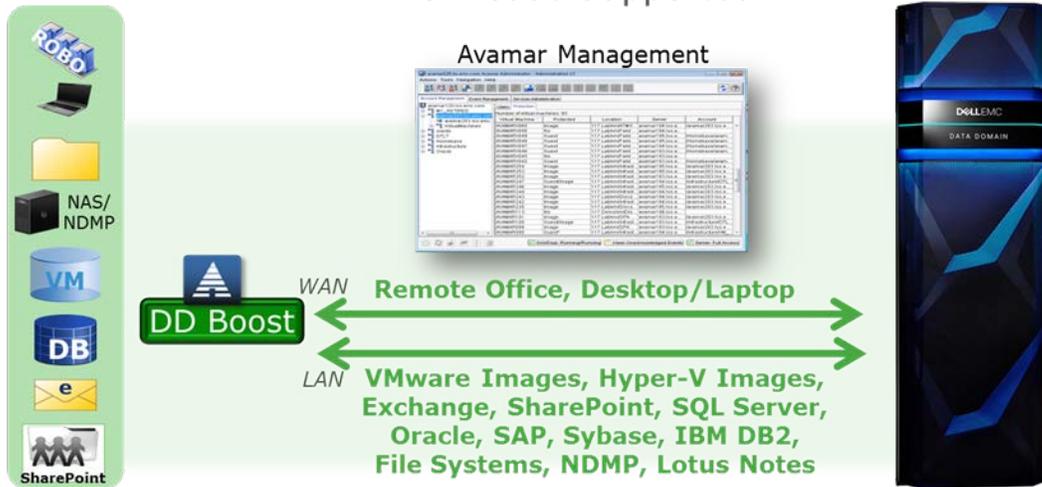


Figure 4: DD Boost for Avamar

DATA DOMAIN BOOST FOR DELL EMC NETWORKER

DD Boost for NetWorker™ includes support at both the Storage Node and all the way to the application client with a feature called Client Direct. The standard NetWorker client now includes DD Boost for all backup workloads including SAP HANA and MEDITECH. This means the up to 99% bandwidth reduction advantage of DD Boost extends all the way from the client to the Data Domain system. It also means that you can now multiply the advantage of distributing some of the deduplication effort to hundreds of clients improving overall backup performance even more. For existing NetWorker customers this means the Client Direct data path bypasses the NetWorker storage nodes, which may currently be overloaded by a high amount of backups. This results in an additional advantage of freeing up storage node resources, which can be used for other purposes. DD Boost for NetWorker also supports DD Boost over Fibre Channel, DD Boost virtual synthetic full backups, and Instant Access to virtual machines stored on Data Domain systems.

DD Boost for NetWorker

- Distribute deduplication to backup server or applications client
- Supported with all NetWorker workloads
- Instant Access for VMware virtual machines

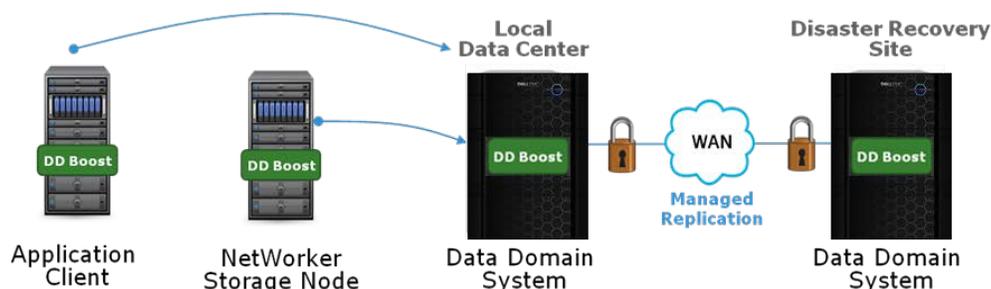


Figure 5: DD Boost for NetWorker

DATA DOMAIN BOOST FOR VERITAS NETBACKUP AND BACKUP EXEC

DD Boost for Veritas NetBackup and Backup Exec leverages their Open Storage (OST) API. OST gives us the opportunity to use a DD Boost plug-in on their media server that provides the DD Boost functionality. With DD Boost, Veritas NetBackup and Backup Exec environments get all the DD Boost advantages noted previously. In addition, DD Boost also supports Veritas AIR for simplified cross-domain disaster recovery, NetBackup virtual synthetic full backups with Data Domain, and NetBackup Accelerator for faster backup of file systems.

DATA DOMAIN BOOST FOR VSPHERE DATA PROTECTION ADVANCED

The Data Domain ecosystem for backup applications now includes VMware's VDP Advanced. DD Boost is part of the VDPA virtual appliance and manages data movement to the Data Domain system. VMware customers now have much greater scalability to store backup data while also having a single deduplication pool for their virtual machines. This becomes much easier for VMware administrators to manage compared to many VDPA virtual appliances, provides increased backup speed, plus the customer gets the highest reliability from the Data Domain Data Invulnerability Architecture. Metadata is stored and managed in the VDPA virtual appliance. VMware administrators can also leverage Data Domain MTree replication to provide bandwidth efficient disaster recovery capabilities.

DATA DOMAIN BOOST FOR HP DATA PROTECTOR

DD Boost for HP Data Protector provides DD Boost functionality over IP at the client level and at the HP Data Protector backup server level. With DD Boost for HP Data Protector, backup administrators get the DD Boost performance, reliability, and bandwidth savings advantages noted previously and can leverage Data Domain MTree replication to provide bandwidth efficient disaster recovery capabilities.

DATA DOMAIN BOOST FOR VEEAM

DD Boost for Veeam integration provides DD Boost functionality over IP or FC at the Veeam client level. With DD Boost for Veeam, backup administrators get the DD Boost performance, reliability, and bandwidth savings advantages noted previously and can leverage Veeam managed replication to provide disaster recovery capabilities.

DATA DOMAIN BOOST FOR COMMVAULT SIMPANA

DD Boost for CommVault Simpana provides DD Boost functionality over IP at the CommVault client level. With DD Boost for CommVault Simpana, backup administrators get the DD Boost performance, reliability, and bandwidth savings advantages noted previously and can leverage Data Domain MTree replication to provide disaster recovery capabilities.

DATA DOMAIN BOOST FOR PIVOTAL GREENPLUM

Pivotal Greenplum™ is a shared nothing, massively parallel processing (MPP) data warehouse Pivotal solution. The Data Computing Appliance (DCA) is the fastest and most cost effective way to leverage Greenplum technology. You can imagine that a large high performance platform like Greenplum creates special data protection challenges and you'd be right. To meet these challenges, Dell EMC has developed DD Boost for Greenplum database that provides all the same performance and efficiency improvements mentioned previously including complete control of backup & replication through the Greenplum database native backup utility. When you have a Greenplum database and a need for serious backup speed and efficient disaster recovery, a Data Domain system is the only solution.

DATA DOMAIN BOOST FOR ENTERPRISE APPLICATIONS

DD Boost for Enterprise Applications leverages DD Boost to give application owners and database administrators direct control over their own backups and restores using native application interfaces they already know and trust. This integration adds DD Boost performance, reliability, and bandwidth efficiency advantages as previously explained. DD Boost for Enterprise Applications also leverages Data Domain MTree replication providing bandwidth efficient disaster recovery capabilities. The following section goes into more detail on how each Enterprise application specifically integrates with Data Domain using native application interfaces.

DD BOOST FOR ENTERPRISE APPLICATIONS

NATIVE APPLICATION INTEGRATION



Figure 6: DD Boost for Enterprise Applications

DATA DOMAIN BOOST FOR ORACLE RMAN

Database backups provide special challenges. Ask any DBA and they will likely tell you they want to perform daily full backups and sometimes more than one full backup each day. But what do you do when full database backups take too long? When a full database backup takes over 12 hours to complete and the database keeps growing, what are you going to do? This is where DD Boost can help. We now have a DD Boost plug-in which can be installed on the Oracle database server achieving performance improvements of up to 50%. That same 12 hour backup over CIFS/NFS might only take 8 hours or less with DD Boost providing the performance advantage that you need to complete backups within your backup window.

In addition, DBAs generally don't like to depend on backup administrators to do their backups and recoveries. DD Boost gives the DBAs the ability to manage their own backups to Data Domain systems using RMAN which they already know and trust. DD Boost with managed file replication also gives them the ability to control replication to a remote Data Domain system for DR with full RMAN catalog awareness of both the local and DR sites.

Data Domain system's logical capacity quota feature also allows the backup administrators to limit the amount of shared Data Domain system capacity that Oracle DBAs can consume to a level which can be agreed upon in advance by both groups. This provides the independence that the Oracle DBAs want and the control that the backup/storage administrators need when using a shared Data Domain storage system.

In addition, using Oracle RMAN directly to Data Domain, customers may avoid expensive backup application licensing charges for each Oracle server being backed up. This makes an investment in Data Domain systems even more cost effective.

DD Boost for Oracle RMAN

- Supported over Ethernet or Fibre Channel
- App Owner can gain control of backup for Oracle RMAN in IP or FC environments
- Choice of MTree or Managed File Replication

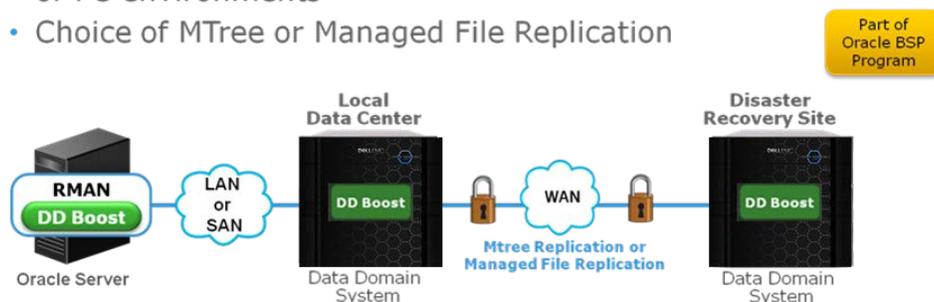


Figure 7: DD Boost for Oracle RMAN

DATA DOMAIN BOOST FOR MICROSOFT SQL

Data Domain Boost for Microsoft SQL provides application centric data protection for Microsoft SQL databases giving the SQL DBA the ability to perform backups and restores directly from SQL Management Studio GUI or CLI to a Data Domain system without the need for an external backup application. This gives the SQL DBAs visibility and control of database backups in addition to the previously mentioned performance, reliability, and bandwidth reduction benefits that are included with DD Boost.

Backups with DD Boost can be performed over Ethernet or Fibre Channel with support for full backups, incremental backups, differential backups, SQL clusters, and SQL Server 2012 Always On groups, and point in time restores. Restores can be performed from a local Data Domain system or a replicated Data Domain system.

DATA DOMAIN BOOST FOR SAP

Data Domain Boost for SAP provides application centric data protection for SAP applications running on Oracle databases giving the SAP administrator the ability to perform backups and restores through the familiar SAP BR*TOOLS GUI or CLI directly to a Data Domain system without the need for an external backup application. This gives the SAP administrator visibility and control of database backups in addition to the previously mentioned performance, reliability, and bandwidth reduction benefits that are included with DD Boost.

Backups with DD Boost can be performed over Ethernet or Fibre Channel with support for full backups, incremental backups, and differential backups. Restores can be performed from a local Data Domain system or a replicated Data Domain system.

DATA DOMAIN BOOST FOR SAP HANA

Data Domain Boost for SAP HANA provides application centric data protection for SAP HANA databases giving the SAP HANA administrator the ability to perform backups and restores through the familiar SAP HANA Studio GUI or the hdbsql CLI directly to a Data Domain system without the need for an external backup application. This gives the SAP HANA administrator visibility and control of database backups in addition to the previously mentioned performance, reliability, and bandwidth reduction benefits that are included with DD Boost.

Backups with DD Boost can be performed over Ethernet or Fibre Channel with support for full backups, online backups, and redo log backups. Restores can be performed from a local Data Domain system or a replicated Data Domain system.

DATA DOMAIN BOOST FOR IBM DB2

Data Domain Boost for IBM DB2 provides application centric data protection for IBM DB2 databases giving the DB2 DBA the ability to perform backups and restores through the IBM Data Studio, DB2 Control Center GUI, or DB2 Command Line Process directly to a Data Domain system without the need for an external backup application. This gives the DBAs visibility and control of database backups in addition to the previously mentioned performance, reliability, and bandwidth reduction benefits that are included with DD Boost.

Backups with DD Boost can be performed over Ethernet or Fibre Channel with support for online or offline backups, full backups, incremental backups, delta backups, and archived log backups. Restores can be performed from a local Data Domain system or a replicated Data Domain system.

DATA DOMAIN BOOST FOR HORTONWORKS AND CLUDERA HADOOP

Big Data use cases are maturing and customers are using Big Data to improve top and bottom line revenues. With this, enterprise readiness and data management needs are becoming increasingly important. Many organizations are working to make their Hadoop & NoSQL environments enterprise ready. Data lakes are the new repository that is becoming a single source of truth.

Data Domain Boost for Hortonworks and Cloudera uses the Hadoop Application Agent which is one agent in a family of agents that make up the DD Boost for Enterprise Applications solution and part of the Data Protection Suite for Applications. DD Boost for Hortonworks and Cloudera provides an effective data protection strategy that addresses these challenges for Hadoop HDFS and HBase Big Data environments leveraging the combined power of DD Boost and BoostFS which will be discussed later in this paper.

As illustrated in Figure 8 below, with DD Boost for Hortonworks and Cloudera, the Hadoop Application Agent installs quickly and easily on the NameNode and provides robust Data Lake protection for HDFS and HBase including the following features:

- Online HDFS backup
- Online HBase table backup

- Backup of existing HDFS snapshots
- Backup retention policies
- Whole or partial HDFS restores
- Whole HBase restores
- Redirected HDFS restores
- Redirected HBase restores
- HDFS restores from replicates
- HBase restores from replicates
- Hadoop High Availability support
- Scriptable command line interface
- Support for multiple configured Data Domain systems
- Kerberos support
- RSA lock box support
- Data Domain deduplication reduces backup storage requirements
- Data Domain Data Invulnerability Architecture data protection
- Can leverage Data Domain replication for Disaster Recovery
- Backup performance better than NFS

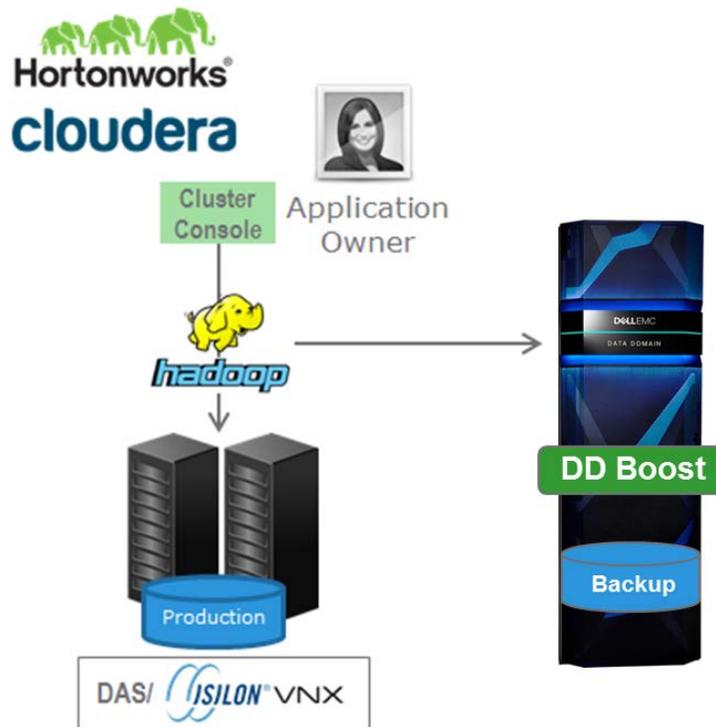


Figure 8: DD Boost for Hortonworks & Cludera Hadoop

As you can see, DD Boost for Hortonworks and Cloudera provides Hadoop (Platform 3) applications versioned backup and recovery data protection for HDFS file system or HBase tables to a Data Domain appliance with the unique advantages of DD Boost that we have talked about throughout this paper.

DATA DOMAIN BOOST ECOSYSTEM

A visual summary of the current DD Boost Ecosystem is illustrated below.

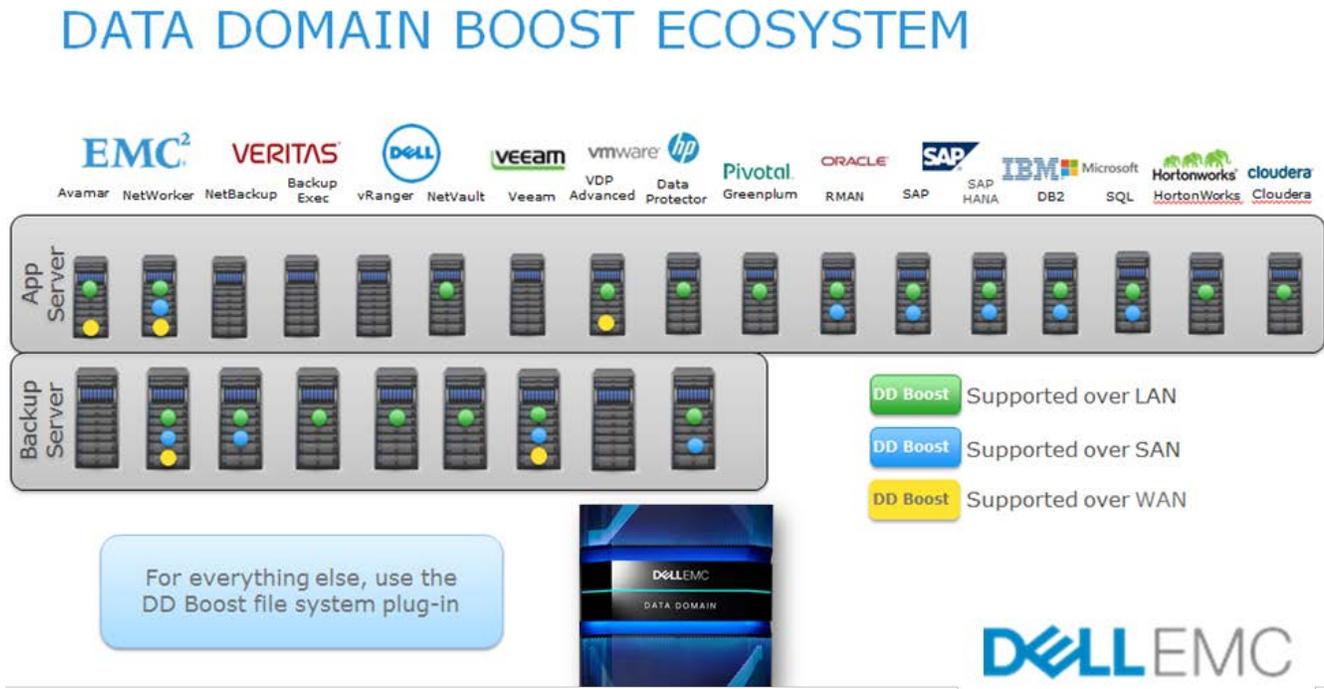


Figure 9: DD Boost Ecosystem

DD BOOST FILE SYSTEM PLUG-IN (BOOSTFS)

Starting in 2010, the DD Boost API program has grown to support over 17 different applications, from backup software to relational databases.

You might be wondering, why did Dell EMC create BoostFS? The answer is simple. There are potentially hundreds of new applications that could benefit from using BoostFS including other backup applications, NoSQL databases, cloud applications or even homegrown customer applications! New applications and workloads now can take advantage of the DD Boost efficiencies without waiting or the cost of advanced application integration. MongoDB, CommVault, and MySQL are only the beginning. In the future Dell EMC will be releasing a tool that will enable developers, partners, and end users to self-validate their workloads for performance, compression, and interoperability with BoostFS then work with Dell EMC for qualification. Bottom line, more customers and more applications will be able to achieve the many benefits of DD Boost through BoostFS. The following paragraphs will explain in a little more detail how BoostFS works.

Introduced in DD OS 6.0, the DD Boost file system plug-in (BoostFS) is a standard filesystem interface that installed on the Linux operating system of your favorite application server. On the client, the filesystem operations conducted on the BoostFS mount point use the Boost protocol to transfer data to and from the Data Domain system. As a result, files and directories created on the mount point are actually stored in the storage-unit on the Data Domain system.

By directly accessing the mount point provided by BoostFS on the client, a third-party data protection application that doesn't have the specific DD Boost API integration can still realize the benefits (e.g. de-duplication, dynamic interface group, TLS encryption) provided by the DD Boost SDK through BoostFS. On the client, users/programs/scripts can access the mount point in the same way they access a local directory.

Through the new DD Boost file system plug-in (BoostFS), DD Boost is now immediately available for new workloads that were previously unavailable and can take advantage of DD Boost benefits. BoostFS can be deployed in minutes, reducing backup windows and storage capacity. Applications using NFS to move data to/from Data Domain can easily switch to BoostFS and improve backup performance.

CONCLUSION

After reading this paper you should have a better understanding how Data Domain Boost can dramatically improve your backup and recovery processes.

To summarize, Data Domain systems with DD Boost will help you:

- Complete backup jobs up to 50% faster within backup windows giving you breathing room for annual data growth.
- Use up to 99% less bandwidth for backup, so you can avoid expensive upgrades by leveraging your existing infrastructure more efficiently.
- Achieve higher backup success with less manual intervention and fewer failed backup jobs to restart.
- Simplify day-to-day backup operations by eliminating the need to manage thousands of tape cartridges.
- Enhance disaster recovery and improve time-to-DR readiness by eliminating all the problems and risks of physical tape based recovery with network efficient, encryption secured replication and total catalog awareness.
- Reduce the workload on your existing backup servers by up to 40%.
- Empower your application owners and DBAs to control their own backups & restores using native utilities for Oracle, SAP, SAP HANA, Microsoft SQL, and IBM DB2.
- Give Hortonworks and Cloudera Hadoop administrators for the first time the capability for true point-in-time backup and recovery of their Hadoop and HBase big data environments.
- Leverage the benefits of DD Boost for third party and Platform 3 applications such as CommVault, MySQL, & MongoDB (e.g. deduplication, dynamic interface groups, TLS encryption) provided by the DD Boost SDK through BoostFS.

If you would like to know more about Dell EMC Data Domain technology differentiation, please refer to other Why Data Domain papers including [Why Data Domain](#), and [Why Data Domain for Oracle](#) and our [Data Domain Data Invulnerability Architecture](#), [Data Domain SISL](#), [Data Domain Replicator](#), and [Data Domain Boost for Oracle RMAN](#) white papers found on EMC.COM. Also, if you have an iPad, please take a look at our [Data Domain Interactive Product Guide](#) eBook on EMC.com. We encourage you to check out [The Protection Continuum](#) blog on this and other Dell EMC data protection and availability topics. You can also visit [the Dell EMC Store](#) to explore Data Domain products.