Deduplication solutions are not all created equal, why data domain?

The Business Value of Data Domain

Why you should take the time to read this paper

- **Speed up your backups** (Achieve up to 68 TB/hr, 1.5 times faster than the closest competitor.)
- **Eliminate the application impact of backups** (Achieve the performance of snapshots and the functionality of full backups with revolutionary primary storage array integration.)
- **Reduce backup costs** (Reduce or eliminate tape infrastructure power, cooling, tape media, and backup application licensing costs.)
- **Improve disaster recovery** (Replace tape-based DR with bandwidth efficient replication improving performance & reliability with simplified DR testing.)
- **Ensure data recoverability** (Dell EMC Data Domain Data Invulnerability Architecture is the industry's best protection for data integrity, which is critical for your storage of last resort.)
- **Simplify backup & recovery operations** (Eliminate tape cartridges and with systems that scale up to 3 PB usable you'll have less storage devices to manage.)
- **Reduce backup & recovery risks** (Eliminate the security risks of using physical tapes for backups with encryption options for data-in-place and data-in-flight.)
- **Save valuable floor space** (Protect 50 PB of logical backups in the footprint of just 2 floor tiles.)
- **Increase backup & recovery service levels** (Maximize success rates via improved performance and reliability.)
- **Facilitate Chargeback & Capacity Planning** (Physical capacity measurement provides the mechanism for chargebacks, trending, capacity planning, and migration planning.)
- **Increase flexibility** (Consolidate backup & archive data and easily adapt to changing requirements over time.)
- **Simplify your purchase decision** (Be confident in your purchase decision by selecting the clear leader in the market with years of proven technology innovation and leadership.)

Table of Contents

Executive summary

Deduplication systems are not all created equal

There is a common misconception that all deduplication systems are created the same and many organizations are now doing their homework prior to making a purchase decision. There are certain key things to look for when you are researching a deduplication storage solution and that is the subject of this paper. Data Domain, powered by Intel® Xeon® processors, is uniquely positioned to deliver you tremendous business value with these important capabilities.

Table stakes or a cut above

All deduplication solutions can reduce your storage and network requirements. However, how efficiently they do it, how fast they do it and whether your critical data can actually be reliably recovered vary greatly. Solutions that are "a cut above" are the ones that don't simply focus on deduplication storage savings, but also provide you the scale, performance and efficient replication you require and prioritize protecting the integrity of your data above all else.

Leaders vs. Followers

Dell EMC continues to lead purpose-built backup appliances with 61.4% total market share – 6x more than the closest competitor – according to IDC. And with over 70,561 Data Domain systems now deployed, we believe the more you know about Data Domain technology, the more you will want to join this group.

Introduction

This paper focuses on Data Domain technology leadership and differentiation and why it matters to you. The purpose of this paper is to explore the technical and financial reasons why Data Domain systems are ideal for backup and archiving in your environment.

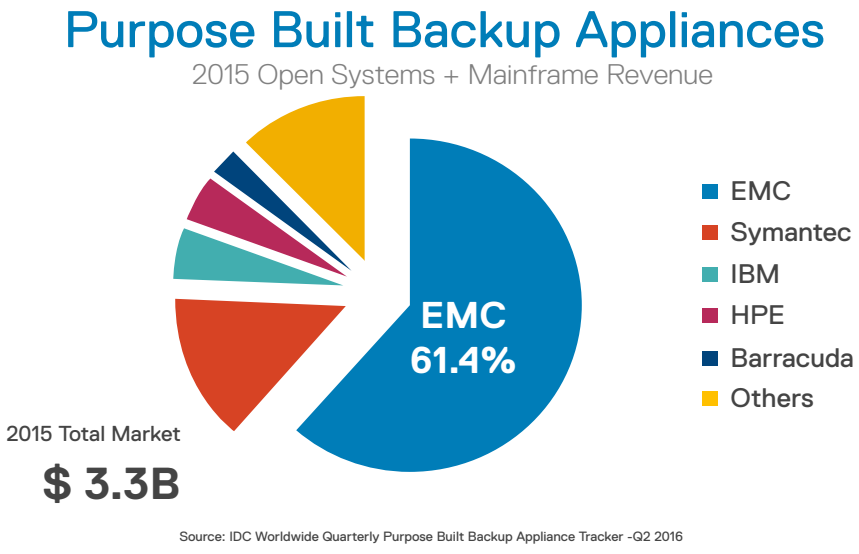
Audience

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 system's differentiating technology and all the unique advantages that it can provide for your backup and archive data.

Why data domain, powered by intel® xeon® processors: technology differentiation & leadership

Data domain data invulnerability architecture

Ensuring data integrity should be priority one for the platform protecting your backup and archive data because they are the storage of last resort. When you try to recover data from this platform, it is likely the only place that data exists. When considering backup and archive solutions, none of the other features and capabilities matter if the data cannot be recovered when it's really needed. No single protection mechanism can provide protection for all the different ways your data can be lost. The Data Invulnerability Architecture includes 4 different protection mechanisms that together provide the industry's best protection for data integrity and recovery.



Source: IDC Worldwide Quarterly Purpose Built Backup Appliance Tracker -Q2 2016

- **End-to-end verification.** The best way to know if the data you are storing is good is to check it after it's been written and compare it against the checksum of the data that was sent. This is done inline when the backup is running so any detected errors can be corrected immediately without having to restart the backup job.

- **Fault avoidance and containment.** One of the most common ways that data gets corrupted is when new data is appended to it, sometimes overwriting previous data. Data Domain systems, powered by Intel® Xeon® processors, avoid this possibility by never appending new data to existing data. The system also includes NVRAM, which protects against data loss in the event of a power failure before all data can be written to disk.

- **Fault detection & healing.** Once you have stored your data correctly, how do you ensure that it stays correct? Over time, bits can flip or become unreadable and disk drives can fail. On all Data Domain systems, an ongoing background process automatically detects and corrects errors on the fly before they become a problem. In addition, RAID 6 protects against a double drive failure or someone removing the wrong drive in the event of a drive failure. Data Domain systems include a global hot spare drive in every shelf. These hot spare drives will automatically take the place of a failed drive and a support call is initiated back to Dell EMC to replace the failed drive.

- **File system recoverability.** Even with all of the above protection, there is always the possibility of a catastrophic failure. For many deduplication systems, these types of failures mean partial or total data loss or may take a week or more to recover from. Since data integrity is the number 1 design priority, Data Domain system uses a self-describing metadata, so we can completely re-build a system in less than 24 hours in the event of this worst-case scenario.

Benefits of data domain data invulnerability architecture

With the Data Invulnerability Architecture, you can reliably recover your critical data and trust that the data will be exactly as you expect it. No other vendor provides this same level of attention to data integrity. Data Domain systems check the data saved to the data that was sent, which ensures your data is stored correctly. In addition, the system takes precautions not to trash existing data by never appending new data to previous data, which ensures your data doesn't get overwritten. Data Domain systems also protect against data loss due to power failures or dual disk drive failures or bit flips with background data scrubbing and on-the-fly error correction, which ensure your data stays recoverable and correct. And finally, unlike most vendors, Data Domain systems leverage self-describing metadata, so the system can rebuild from scratch in a reasonable timeframe, to ensure you're up and running as quickly as possible. This commitment to ensuring data integrity should give you confidence to trust Data Domain systems to protect your data better than anyone else.

Data domain stream informed segment layout (sisl™)

The foundation for Data Domain's industry leading performance is the Stream-Informed Segment Layout (SISL) scaling architecture. SISL 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. SISL means Data Domain systems do not rely on increasing the number of disks to increase performance and therefore are not spindle-bound like other deduplication platforms. This is why Data Domain systems have dramatic increases in performance with each successive generation of Intel processors – every time Intel processors get faster, Data Domain systems get faster

Benefits of SISL

There are 2 important benefits from SISL – faster backups and investment protection. Most importantly, since Data Domain systems are the fastest in the industry, they will help you meet tight backup windows in the face of exploding data growth. Secondly, because Data Domain systems performance increases with Intel performance, it follows Moore's Law. This means that future Data Domain systems will continue to realize dramatic improvements in speed and scalability as future CPUs are used in new Data Domain systems. As new technology is introduced, many of our systems enable you to replace the controller with a next generation model while leaving all the backup data in-place. This investment protection ensures you can dramatically improve backup performance and scalability without disrupting operations.

Dell EMC data domain boost™ software

Dell EMC Data Domain Boost software distributes parts of the deduplication process to the backup server(s) or application client(s), leaving the Data Domain system, powered by Intel® Xeon® processors, 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. The larger the backup shop the more significant this distribution is. A backup shop with five or more backup servers, for example, would have five backup server resources each doing some of the deduplication effort with DD Boost. Without DD Boost, the entire deduplication effort is being performed 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 down to the client and in these cases, the distribution benefit isn't five (backup servers) to 1, but could be hundreds or thousands (clients) to 1.

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 the DD Boost File System Plug-In, or BoostFS. On the client, users/programs/scripts can access the mount point in the same way they access a local directory.

Benefits of data domain boost

DD Boost speeds up backups by 50% without changing your existing backup servers and infrastructure. Doesn't that sound great … speed up your backups using the same exact hardware? A single controller DD9800 has performance that is 1.5 times faster than the closest competitor achieving backup speeds up to 68 TB/hr!

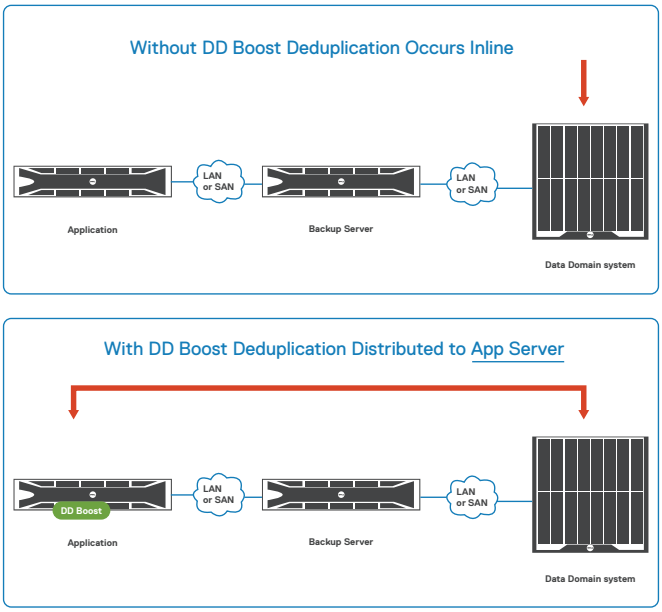
With DD Boost, only unique data has to be sent from the backup server or client to the Data Domain system. This means up to 99% less data has to be moved across the network - even for full backups - allowing more efficient use of your existing resources. For applications that DD Boost can be leveraged at the client (Dell EMC NetWorker™, Dell EMC Avamar™, Oracle RMAN, NetVault), this bandwidth reduction spans the entire backup path all the way from the client to the Data Domain system.

In addition to increased performance, there's another advantage of distributing the deduplication process that may not be so intuitive. Specifically, DD Boost actually reduces CPU utilization on the backup server or client even though it's executing parts of the deduplication process. As it turns out, the CPU cycles required to execute these parts of the deduplication process are actually less than what it takes to push full backups over the network. Aha, now that's pretty cool, huh? With DD Boost, backups run faster, you use less bandwidth and you reduce the workload for your backup server or client. Wow! But wait, there's more.

DD Boost also means you don't have to manage thousands of physical or virtual tape cartridges greatly simplifying your day to day production and disaster recovery operations and reducing the time, effort and costs associated with handling and managing tape cartridges.

Even with deduplication, managing replication can be difficult and DR testing can be cumbersome. DD Boost with managed file replication changes this by providing the backup application visibility and control over Data Domain replication. This gives the backup application total catalog awareness of all local copies and any copies that have been replicated to other sites and increased confidence in your disaster recoverability.

Finally, 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, which improves the reliability of your



backups and eliminates the need to manage mount points. This also ensures your backups continue to run even if you lose a path resulting in higher backup completion success percentages and less effort spent on re-running failed backup jobs.

DD Boost for Enterprise Applications also gives application owners the control and visibility that they've always wanted in addition to all the other DD Boost benefits.

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.

Variable-length segmentation

Data Domain systems, powered by Intel® Xeon® processors, use variable-length segmentation to break up data streams for optimal deduplication rates. Specifically, as a Data Domain system ingests data, it intelligently breaks up the stream based on the natural structure of the data. Then, the system will determine if each segment is unique before compressing and storing it. By doing this, the system usually finds the same logical segment break points that it found previously, which results in identifying more duplicate segments enabling higher deduplication ratios. In comparison, vendors who use fixed length deduplication are less likely to find duplicate segments as their segmentation is based on a predetermined size. Some vendors will even try to trick you and call it "variable" when it really means you can select which "fixed" length value you want to use.

Benefits of variable-length segmentation

Given the same real world data set, variable-length deduplication will always produce higher deduplication ratios. The significance of this cannot be overstated. Variable-length segmentation equates to higher deduplication ratios, which means you need less storage to protect your data. In addition, this also enables more effective scalability within a single pool, which means you'll have fewer devices to manage. Finally, the higher deduplication ratios that variable-length deduplication enables means you will have less data to replicate and therefore require less WAN bandwidth and less storage at your DR site. All this adds up to significantly less complexity & cost!

Inline vs. Post process deduplication

Data Domain systems perform deduplication inline, as the backup stream comes into the system and only stores unique elements on disk. Unlike post-process deduplication, they do not have to store data on disk first and then deduplicate it afterwards as a separate process.

Benefits of data domain inline deduplication

Inline deduplication means Data Domain systems do not have to include additional storage capacity as a landing zone for backup data so that it can be deduplicated later. This means less storage, less cost and less footprint in your data center. After all, isn't a smaller storage footprint one of the main reasons for deduplication?

Massive scalability

Data Domain systems, powered by Intel® Xeon® processors, offer tremendous scalability with up to 1 PB usable capacity in the active tier of a single DD9800 system that fits in the footprint of just 2 floor tiles and can protect up to 1 billion small files. The DD9800 can protect up to 50 PB of logical backup data all in a single deduplication pool. Systems can start with as little as a few disks and scale up to 24 shelves for the active tier. Data Domain systems scale seamlessly without disrupting operations by simply adding additional shelves on-the-fly while the system is running. Leveraging the DD Cloud Tier option, a second tier of storage can be added to some models for long-term backup retention, which provides further scalability up to 3 PB of total usable capacity or up to 150 PB of total logical storage.

Mid-range and high-end Data Domain systems can leverage the DS60 dense shelf option, which can provide up to 190TB of usable capacity in only 5U of rack space! Better yet, all Data Domain systems come with a shelf migration capability that enables older drive/lower density shelves to be replaced with newer drive/higher density shelves while the system continues running with minimal performance impact.

Benefits of data domain scalability

Massive scalability means you will have fewer devices to manage, require less infrastructure and achieve higher deduplication ratios because there can be more data within a single deduplication pool. Other vendors offer smaller scalability that may require you to deploy many systems, meaning multiple deduplication pools and more complexity, which results in a lower overall deduplication efficiency. With Data Domain systems, shelves can be added while the system is running providing additional scale without disruption. This massive scalability enables Data Domain systems to provide the capacity required for efficient consolidation for backup and archive data. Data Domain Cloud Tier and extended retention options also allows you to eliminate the problems, risks and expenses associated with using physical tape for long term backup retention.

The DS60 dense shelf option provides maximum data center footprint cost efficiency. The shelf migration utility simplifies

operations over time, maintains overall Data Domain system investments and maximizes data protection availability.

Data domain for disaster recovery

With tape-based disaster recovery, there are many risks and recovery dangers including damaged tape media, lost tape media, tape drive hardware failures, tape storage recovery delays and frequently a limitation on the number and speed of tape drives available at the recovery site. While most disk based storage platforms offer replication, without deduplication, those alternatives are not practical or cost effective for disaster recovery.

Data Domain systems with Data Domain 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. Data Domain's efficient inline variable length deduplication becomes the enabling technology for a cost effective "tapeless" disaster recovery approach. Specifically, with Data Domain Replicator, Data Domain systems only replicate unique data to the remote site and begin replication while backups are still in process.

Data Domain systems also make regular DR testing easier and faster through a unique snapshot capability called fastcopy. Fastcopy is a metadata read/write snapshot that can be created in less than 5 minutes to be used for DR testing.

Benefits of data domain for disaster recovery

The first and most obvious benefit of Data Domain Replicator is 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 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 read.

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 even have virtual tape cartridges to worry about. Even better, with DD Boost and managed file replication, your backup catalog will already be fully aware of all replicated copies available at your DR site.

Data Domain fastcopy provides for quicker and easier DR testing without impacting production replication from the primary site. Fastcopy uses almost no additional storage capacity in the Data Domain system at the DR site because it uses metadata pointers to existing deduplicated data. When DR testing is complete, that fastcopy snapshot can be safely and easily deleted.

Physical capacity measurement

Data Domain systems, powered by Intel® Xeon® processors, can provide reports on physical capacity used by files, MTree, or by Tenant to facilitate chargeback billing, capacity planning, migration planning and provide insight into top Data Domain protection storage consumers. These reports can be run on-demand or batch.

Benefits of data domain physical capacity measurement

Data Domain physical capacity measurement provides an easy way for Enterprise Customers or Service Providers to measure consumption of Data Domain physical capacity usage providing a chargeback methodology and capacity usage information which can be used for capacity planning or migration planning.

Secure multi-tenancy

For large enterprise customers and for service providers, Data Domain systems provide secure multi-tenancy capabilities for secure data isolation, management and reporting by internal business units or departments, or individual customers.

Benefits of secure multi-tenancy

Secure multi-tenancy provides the ability to share a physical Data Domain system while providing logical data isolation by tenants and administrative management and reporting isolation by tenant. This improves overall cost efficiency by enabling greater storage consolidation, simplifies on-going management and provides the basis for chargeback for protection storage.

Oracle optimized deduplication

Deduplication efficiency depends on backup streams looking very much the same from day-to-day. This is a result of typically starting backups at the same point and going in the same order each time. Data Domain uses the most efficient variable length segmentation approach to determine logical places to segment the incoming backup stream. The result of starting at the same place, going in the same order and efficient variable length segmentation is achieving high deduplication ratios up to 30:1 with typical enterprise data and retention periods.

In the physical tape backup world, multiplexing is very common. Multiplexing means sending multiple backup streams to a single target device mixing the backup data in order to keep the tape buffer full for the physical tape drive so that it functions as fast

as it can. Multiplexing is typically enabled by a setting in the backup application.

It is also very common to backup Oracle databases using multiple channels in order to improve overall backup performance. When the Oracle filesperset value is set greater than 1, multiple channels are used. This is, in fact, another way to multiplex backups.

Unfortunately the multiplexing methods mentioned above have a negative impact on deduplication efficiency because it varies the backup stream from day-to-day. When the backup data is mixed into a common stream, even though they may start in the same place, the data won't be in the same order every day. The result is significantly less deduplication efficiency. This is true for any vendor's deduplication solution. For this reason, the standard deduplication best practice is to turn off multiplexing and set Oracle filesperset = 1. Until now, you have been forced to choose between high performance and high deduplication efficiency. With Data Domain Oracle optimized deduplication technology, you can use multiplexing and still achieve high deduplication efficiency for Data Domain systems dedicated to Oracle DB backups.

Benefits of oracle optimized deduplication

Data Domain's Oracle optimized deduplication technology means you no longer have to choose between maximizing speed or maximizing deduplication efficiency, you can have both at the same time for systems dedicated to Oracle DB backups. Multiple Oracle channels maximize database backup performance. Higher deduplication ratios mean less storage used, less bandwidth used and in the end, less cost and complexity.

Protectpoint: the performance of snapshots with functionality of backups

How do you take industry leading protection storage and make it even better? Dell EMC has integrated its best of breed primary storage with its best of breed Data Domain protection storage creating a revolutionary new backup solution called ProtectPoint. Unlike traditional backup, ProtectPoint will only pause the application to mark the point in time for an application consistent backup and then the application can quickly return to normal operations. ProtectPoint sends data directly from the primary storage array to the Data Domain system over Fibre Channel. Leveraging change block tracking technology, only the data that has changed since the last full backup is sent directly from the primary storage array to the Data Domain system, powered by Intel® Xeon® processors.

ProtectPoint is integrated with Oracle RMAN including support for RAC, SAP with Oracle using BR Tools and IBM DB2 for open systems using IBM Data Studio. This gives Oracle, SAP and IBM database administrators the ability to leverage ProtectPoint from native utilities that they are already familiar with. Dell EMC NetWorker has been integrated with ProtectPoint providing the NetWorker backup administrator visibility into all ProtectPoint data protection activity. In addition, ProtectPoint also supports file system backups for Microsoft Windows as well as Unix and Linux.

Benefits of data domain with protectpoint

Because ProtectPoint leverages change block tracking technology, only the data that has changed gets sent to the Data Domain system, which greatly reduces the amount of data that needs to be moved. Because the data movement is directly from the primary storage array to the Data Domain system, ProtectPoint also eliminates the application performance impact of traditional backups, which enables more cost effective retention and allows for faster and more frequent full backups with less cost and complexity. ProtectPoint provides the performance benefits of snapshots with the functionality benefits of full backups. This is particularly valuable for the protection of very large datasets. ProtectPoint's database integration provides all these advantages using tools that are already familiar to the application owners. And finally, for NetWorker customers, ProtectPoint integration provides corporate backup administrators visibility & awareness of all ProtectPoint activity.

Flexibility

Most of us don't like to be locked into one way of doing things when we purchase a solution. Data Domain systems provide investment protection with the flexibility to grow with you as your backup requirements change over time. All Data Domain systems include 1Gig Ethernet ports that allow you to quickly and simply perform backups with CIFS, NFS as a NAS target. You also have the option of adding one or more dual port 8GB or 16GB Fibre Channel HBAs to connect your Data Domain system into an existing FC infrastructure and in many situations will be able to leverage DD Boost and eliminate the need to manage tape cartridges. Data Domain systems also support NPIV FC port virtualization. You also have the option to add one or more quad port 1 Gig or dual/quad port 10 Gig Ethernet NICs into your Data Domain system to support additional bandwidth over Ethernet.

In addition, you can perform backups over Fibre Channel and Ethernet at the same time. All data in the Data Domain system is part of the same deduplication pool regardless of how the data gets into the system. You can also use Data Domain as the target from multiple backup and/or archiving applications at the same time and set logical quotas for each workload. In addition, many models have a cost effective controller upgrade option leaving your existing storage and backup data in place.

Data Domain supports all the leading Open Systems backup applications including Dell EMC Avamar and Dell EMC

NetWorker and can also be used with IBM i systems and IBM zSeries mainframe systems.

The Data Domain Cloud Tier option supports native Data Domain long term backup retention to private or public clouds

Benefits of data domain system flexibility

Data Domain's flexible connectivity means you can easily drop a system into an existing infrastructure with almost no change to your existing backup processes and be up and running in no time. Then, you can add functionality over time as your requirements and environment changes. For example, to minimize change, you might start out by installing a new Data Domain system using DD VTL software over Fibre Channel. Later, you could install a multi-port Ethernet NIC and eliminate using tape cartridges, or a multi-port Fiber Channel HBA card and with DD Boost you can eliminate using tape cartridges. And NPIV FC port virtualization support provides maximum FC connectivity and efficiency ensuring Data Domain can easily be inserted into all environments with maximum FC throughput.

You can also share a Data Domain system between multiple backup and archiving applications at the same time. This is ideal if you have multiple different backup applications (possibly due to acquisitions) and you want to migrate from one backup application to another over time without disruption. In addition, by consolidating backup and archive data on a single system, you can eliminate silos of storage and a lower TCO. Data Domain systems also provide investment protection by supporting your non-Dell EMC backup applications today, with the option to upgrade later to Dell EMC Avamar or Dell EMC NetWorker for optimal performance and end-to-end integration.

And finally, many Data Domain systems allow non-destructive upgrades by replacing the controller with a newer model and keeping all existing backup data in place on your existing storage shelves. Since no data migration is required, this can be a cost effective way to upgrade your performance and scalability while leveraging your initial investment in storage shelves.

Data Domain's Cloud Tier option for native long term backup retention lets customers choose whether to leverage public cloud object storage for long term deduplicated backup retention or send it to their own private cloud target such as Dell EMC Elastic Cloud Storage

Consolidation platform for backup and archive

Data Domain systems, powered by Intel® Xeon® processors, are not only industry leading purpose-built backup appliances, but are also an ideal platform for archive data. Data Domain systems support many leading archive applications such as Dell EMC SourceOne, Symantec Enterprise Vault and IBM InfoSphere Optim Archive. By integrating with these file, email, SharePoint and database archiving applications, Data Domain systems provide efficient archive storage through deduplication.

Data Domain Retention Lock Governance edition provides file level locking capability. The Data Domain Operating System includes data shredding capability. Data Domain Retention Lock Compliance edition, provides secure data retention for file and email archive data that meets US and International standards including SEC 17a-4(f).

With DD Retention Lock, all Data Domain systems can simultaneously support governance and compliance archive data sets. This enables you to set different retention periods for different classes (governance and compliance) of archive data.

In addition, Data Domain systems also offers Data Domain Cloud Tier or Data Domain Extended Retention software options that can eliminate using problematic physical tape for long-term backup retention. DD Cloud Tier enables up to 150 PB of logical capacity in a single system for long-term backup retention. This software option is available for new or existing DD860, DD990, DD4200, DD4500, DD7200, DD9500, DD6800, DD9300 and DD9800 systems. The Data Domain Cloud Tier option allows customers to send deduplicated long term backup retention data to a public or private cloud object storage target.

Benefits of data domain as a consolidation platform

Data Domain systems are ideal for consolidating backup and archive data to reduce overall TCO by eliminating silos of storage. This enables both workloads to benefit from the deduplication storage and replication efficiencies and means you'll only have one device to manage.

In addition, DD Cloud Tier and DD Extended Retention provide cost effective alternatives to physical tape for long-term backup retention eliminating the risks and costs of handling, storing and managing thousands of tape cartridges. DD Cloud Tier offers native deduplicated long term retention to private or public clouds. The minimal day-to-day attention that Data Domain requires makes it a perfect consolidation platform.

Hardened security

Data Domain systems have a number of important security features that protect data being stored and replicated. Role Based Access Control provides access to Data Domain resources based on what the individual needs to do. These roles include administrator, backup operator, user, data access and Security Officer.

Data Domain Replicator enables safe and network-efficient replication by encrypting all or a portion of the data to be replicated. Data is deduplicated as it is being written to the Data Domain system and DD Replicator preserves this

deduplication, thereby reducing the network utilization.

The Data Domain Encryption software option provides the industry's first encryption of data-at-rest on deduplication storage, enabling organizations to enhance the security of their data. Data Domain systems have local key management capability.

Benefits of data domain hardened security

Role Based Access Control provides security protection by limiting action that can be taken by users either intentionally or unintentionally based on what their role is.

Optional encryption of data-at-rest provides an extra layer of data security, if required, for data that sits on the Data Domain system. Optional encryption of data-in-flight is available to protect data being replicated from one Data Domain system to another while preserving the deduplication bandwidth efficiencies of Data Domain Replicator.

Data Domain includes basic encryption key management capabilities and optional integration with RSA Data Protection Manager, which can manage encryption keys for Data Domain and other systems. This integration option with RSA gives customers the ability to manage all their encryption keys through a single mechanism.

Data domain virtual edition

Data Domain Virtual Edition (DD VE) leverages the power of the Data Domain Operating System to deliver software-defined protection storage. DD VE is fast and simple to download, deploy and configure and can be up and running in minutes. DD VE runs in VMware vSphere ESXi 5.5 and 6.0 and supports VMware vSphere High Availability and Fault Tolerance to meet the availability needs of customers. It also supports VMware Distributed Resource Scheduler (DRS), allowing VMware to balance workloads for optimal performance. DD VE also runs in Microsoft Hyper-V. DD VE provides customers with the benefits of the world's most trusted protection storage and the simplicity, flexibility and efficiency of a software-defined solution.

DD VE maintains the core Data Domain features that differentiate it as the industry-leading protection storage. This includes high-speed, variable length deduplication for a 10 – 30x reduction in storage requirements, unparalleled data integrity to ensure reliable recovery and seamless integration with leading backup and archiving applications. DD VE also comes with DD Boost, which speeds backups by up to 50%, DD Encryption for enhanced security of data and DD Replicator, which enables network efficient replication for faster time-to-DR readiness.

Benefits of data domain virtual edition

A single DD VE instance can scale from .5 TB to 96 TBs. Capacity can easily be moved between virtual systems and/or locations and can be purchased in 1 TB increments allowing you to grow capacity as the business demands it. This provides customers with tremendous deployment flexibility. DD VE enables customers to gain the benefits of the world's most trusted protection storage with the agility, flexibility and efficiency of a software-defined solution.

Data domain high availability option

The Data Domain High Availability option (HA) ensures the operational continuity of backup, archive and recovery of data to minimize downtime for users and processes. The HA option is available for the DD9800, DD9500, DD9300 and the DD6800 systems. The active/passive configuration attaches two Data Domain controllers to a shared storage pool, with one handling data ingestion and the other on standby. The high availability interlink card mirrors the state of the active node and NVRAM content between controllers. During unplanned system downtime such as a sudden hardware failure, failover activates. Backup jobs will pause on the active controller and failover to the standby node, where they can resume operations in just minutes. HA on the Data Domain system fails over automatically for DD Boost and NFS protocols, allowing users to effortlessly maximize uptime in the face of an unexpected interruption.

With the HA option, Data Domain users gain the ability to upgrade the Data Domain Operating System (DD OS) without having to take a system offline. When the process is initiated, first the standby controller will be upgraded while the active is still operating. Once the first upgrade is completed, operations will failover to the updated controller and the upgrade will begin on the second system.

Benefits of data domain high availability option

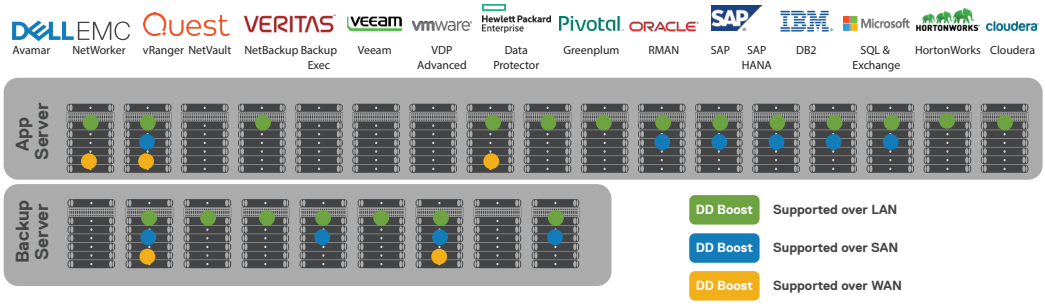
The Data Domain High Availability option enables users to achieve greater operational resiliency in the face of unexpected failures, with failover enabled between two Data Domain controllers. Enterprise customers facing increasing demands for business continuity will rest easier with a second Data Domain controller on standby at all times. The primary benefits of high availability are:

- Dramatically minimize downtime: get back up and running in minutes with HA, enabling business continuity in the face of sudden hardware failures
- Faster system upgrades: upgrade controllers without having to take a system offline

DD Boost solution integration

Dell EMC continues to innovate and leverage Data Domain with DD Boost in many ways to benefit our customers. Integration with leading backup and enterprise applications continues to grow with ever increasing opportunities for customers to take advantage of this powerful technology.

Data Domain Boost Ecosystem



For everything else , use the DD Boost file system plug -in

DD Boost Integration:

- Dell EMC NetWorker
- Dell EMC Avamar
- VMware vSphere Data Protection
- Veritas NetBackup
- Veritas Backup Exec
- NetVault
- vRanger
- HP Data Protector
- Veeam
- CommVault Simpana
- Pivotal Greenplum

DD Boost for Enterprise Applications:

- Oracle RMAN
- Microsoft SQL
- SAP
- SAP HANA
- IBM DB2
- Hortonworks Hadoop
- Cloudera Hadoop
- Mongo DB
- MySQL

Conclusion

After reading this paper you should have a better understanding how Dell EMC Data Domain systems, powered by Intel® Xeon® processors, can dramatically improve your backup, recovery, archive and long term retention processes.

To summarize, Dell EMC Data Domain will help you:

- Complete your backups quickly and give you some breathing room to handle annual data growth within your backup windows.
- Improve protection and enhance recovery of large database backups by dramatically reducing backup times while reducing application impact.
- Reduce the strain on your network infrastructure caused by backups and defer the cost of possible infrastructure upgrades.
- Free up valuable data center floor space by eliminating large physical tape libraries and tape storage.
- Achieve higher backup job completion success with less manual intervention.
- Reduce your ongoing backup and recovery costs related to power, cooling, tape management and backup licensing.
- Be confident in the recoverability of critical data when you really need it.
- Optimize Oracle database backups leveraging the speed of multiple channels while maintaining high deduplication efficiency.
- Have happier Oracle database administrators by giving them dramatically faster daily full database backups with full catalog visibility and control over their own backups and recovery.
- Simplify day-to-day backup operations by eliminating the need to manage thousands of tape cartridges.
- Enhance disaster recovery and improve RTOs by eliminating all the problems and risks of a physical tape based recovery with bandwidth efficient and encryption secured replication.
- Reduce the time, effort and costs for annual disaster recovery testing eliminating the need to transport and manage physical tape.
- Enhance cost efficiencies and facilitate chargeback for protection storage by internal business units or customers and Service Providers leveraging secure multi-tenancy.
- Empower Hadoop admins to do their own backup and recovery on Hortonworks and Cloudera's Hadoop distribution.
- 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 Data Domain technology, please refer to our Data Domain Data Invulnerability Architecture, Data Domain SISL, Data Domain Replicator, Dell EMC ProtectPoint, and Data Domain Boost for Oracle RMAN white papers. Please join us on The Core blog to discuss this and other EMC data protection and availability topics. You can also visit the Dell EMC Store to explore Data Domain products.