



EMC NetWorker Rounds Out Deduplication Support with EMC Data Domain Boost

Analyst: Michael Fisch

Management Summary

Now and then, a solution comes along that is artful enough to make one plus one equal three. EMC *Data Domain Boost* enables such a solution. When used in conjunction with a supported backup application – in this case, its most recent integration, EMC *NetWorker – DD Boost* simultaneously speeds up the backup process, reduces the impact on the backup server, and centralizes and simplifies management of deduplication backups within the *NetWorker Management Console*. All of this is accomplished through software, so existing Data Domain customers will not require any additional hardware.

The NetWorker integration with DD Boost, following its integration with EMC Avamar, has created a unified and centrally-managed solution set for traditional and deduplication backup. Tethering Avamar and Data Domain to NetWorker enables this backup application to offer an evolutionary means for enterprises to adopt next-generation technologies. Read on for details about why we believe DD Boost and its integration with NetWorker will be a hit.

NetWorker Gets Data Domain Boost

EMC's integration of NetWorker and Data Domain Boost has resulted in a complete software and hardware solution to enable a much simpler, more predictable, and more supportable solution.

EMC Data Domain

EMC Data Domain is the most widely adopted deduplication storage system in the industry. It is an inline deduplication appliance that easily integrates into a backup environment by connecting to the backup application as a file server (CIFS, NFS) over IP, as a virtual tape library over Fibre Channel, or now through the backup-optimized DD Boost data access method. Data Domain delivers high-performance throughput and typically 10-to-30-times data reduction. These systems also offer network-efficient replication of deduplicated data to a disaster recovery site.

EMC NetWorker

EMC NetWorker is an enterprise backup application with a large customer base. EMC has evolved NetWorker to become a unified platform for traditional and next-generation backup and recovery capabilities. NetWorker provides centralized backup and recovery for a wide range of heterogeneous environments, both physical and virtual. Its features include support for SAN, NAS, and DAS storage, a variety of applications and operating systems, off-host "hot" backup while applications are in use, and 256-bit AES file encryption and user authentication for data security. Today, NetWorker also supports data deduplication through cohesive integration with Avamar client-based and, now, Data Domain deduplication storage systems.

IN THIS ISSUE

➤ NetWorker Gets Data Domain Boost	1
➤ EMC NetWorker Plus Data Domain and Avamar	3
➤ Conclusion	4

Data Domain Boost

DD Boost is like a turbocharger for Data Domain deduplication storage systems. It is a software option that performs the hat trick of simultaneously speeding up the backup process, reducing the impact on the backup server, and simplifying deduplication management within the backup application. Since this is accomplished through software only, no additional hardware is necessary for existing Data Domain customers.

Distributed Deduplication for Faster Performance

According to EMC's figures, DD Boost delivers significantly faster backup performance:

- Up to 50% faster backups,
- 20% to 40% less impact on the backup server, and
- 80% to 99% less backup traffic on the LAN.

So how does DD Boost achieve faster performance and less impact? You could say it is an outside-the-box solution, in a literal sense, because DD Boost offloads parts of the deduplication processing to the backup server while performing the remainder on the Data Domain system. (Without DD Boost, the Data Domain system handles all of the deduplication processing.) A NetWorker storage node integrated with DD Boost performs parts of the deduplication process by breaking data into variable-length segments. The storage node then checks with the Data Domain system to determine if the segment is unique and then compresses and sends any unique data segments over the network. The Data Domain system on the receiving end completes the deduplication process and then writes the unique, compressed data to disk.

DD Boost also optimizes the communication between the storage node and Data Domain system. In backup environments, disk storage devices typically present themselves either as virtual tape libraries (VTLs) or as file servers running *CIFS* or *NFS*. Data throughput is subject to the overhead and limitations of those protocols. However, DD Boost is optimized specifically for fast throughput of backup data.

DD Boost reduces the overall resource utilization on the NetWorker storage node because the added deduplication processing is more than offset by a reduction in I/O pro-

cessing. It first does a comparison to identify redundancies, so it is able to send much less data over the network. Data movement, in this case, is the heavy lifting. Therefore, the net effect is to reduce CPU utilization on the backup server and complete backup jobs faster.

EMC integrated the DD Boost software component directly into the NetWorker storage node. This approach avoids the overhead of going through an API layer on the backup server. Thus, the integration is maximally streamlined for the best performance and functionality.

DD Boost deduplication is inline, as it normally is for Data Domain systems. Inline deduplication means data is processed in real time as it is backed up, as opposed to using post-processing techniques that deduplicate periodically and after backup jobs are completed. In fact, DD Boost's distributed deduplication is only possible with an inline approach because it needs real-time processing. Inline deduplication uses less storage capacity, since it does not require a disk buffer for un-deduplicated data. It also allows for immediate remote replication of deduplicated data for better protection and DR readiness.

Managing Data Domain within NetWorker Management Console

DD Boost also enables Data Domain system management from the NetWorker Management Console, giving administrators a single point of management for deduplication backup and recovery.

Discovery and Configuration Wizards

With DD Boost integration, the NetWorker Management Console provides a wizard that walks users through the process of setting up Data Domain systems for deduplication backup. NetWorker automatically discovers the systems and assigns them their own "Data Domain" device type – which is different from an AFTD (Advanced File Type Device) or VTL.

Once a Data Domain device is configured, it can serve as a target for client backups, and administrators can set policies for deduplication backup in NetWorker, just like other target devices.

Clone Controlled Replication

With DD Boost, the cloning workflow in NetWorker can manage and control the replication of Data Domain systems. Cloning is the NetWorker feature for creating copies of data

sets. Traditionally, it was used for copying backup tapes, though it applies to all backup media – disk and tape. The cloning workflow allows users to schedule automatic copies and define retention periods. It catalogues all available copies.

On the storage end, Data Domain Replicator is the software option that replicates data between and among local and remote Data Domain systems. Replication is used for DR, branch office data protection, multi-site backup consolidation, and long-term archiving. Since DD Replicator sends only unique data segments across the network, it requires significantly less bandwidth than other asynchronous replication technologies.

Cloning controlled replication allows NetWorker to be the single point of management for deduplicated replication and traditional full backup copies. Users can establish schedules at the times chosen as most appropriate by the administrator. They can configure one-to-one as well as one-to-many and multi-site replication. NetWorker tracks replicated copies in its clone catalog. It can manage separate retention policies for local and remote copies. Furthermore, if a NetWorker storage node is available at the remote site, the Data Domain replica can be used as a source for creating full tape copies from the storage node for long-term, offsite archiving – without the burden of physical transportation.

Monitoring and Data Collection Report

NetWorker has the ability to monitor and report on Data Domain. Within the NetWorker Management Console are seven different reports about Data Domain systems and their associated backups. The monthly and daily summaries are especially useful for administrators to see an overview and status of the backup environment. NetWorker will display deduplication ratios at both the saveset and system level directly through the NetWorker Management Console.

For monitoring and system maintenance, NetWorker tracks up to 25 SNMP traps for Data Domain systems. The traps automatically alert the administrators when certain device-specific events occur, such as:

- *File Space Alarm* – Unused capacity is running low
- *NVRAM Failing* – Electronic memory needs to be replaced

- *Power Supply Failed* – Power supply needs to be replaced
- *No Spare Disk* – An extra disk needs to be added in case of a RAID rebuild

These monitoring and reporting features allow NetWorker to be a single point of management and device maintenance for backup environments that include Data Domain deduplication storage systems. This adds to NetWorker's capability as a unified platform for enterprise backup and recovery.

Additional Details

DD Boost is supported only on IP networks. Enterprises that want to run Data Domain over a Fibre Channel SAN alternatively can use the VTL software option.

EMC NetWorker Plus Data Domain and Avamar

The NetWorker integration with Data Domain Boost, following its tight integration with EMC Avamar¹, has created a unified and well-rounded solution set for traditional and deduplication backup. NetWorker² itself is a top-tier enterprise backup application with a substantial installed base. Avamar is the industry's leading client deduplication backup software. Data Domain is the industry's leading deduplication storage system. The three combined with a single point of management, scheduling, and reporting in the NetWorker Management Console form a great set of capabilities for disk-based backup and recovery. NetWorker becomes the bridge and an evolutionary path forward from traditional to next-generation backup technologies.

Moreover, disk-based backup is the most significant trend in backup and recovery over the last decade, and data deduplication is the key enabling technology. Disk is much faster than tape for both backup and restore, and enterprises have adopted it en masse as a way to handle increasing data protection requirements. Data deduplication has facilitated this adoption because it reduces backup data by 10

¹ See [The Clipper Group Navigator](#) dated March 30, 2010, entitled *EMC NetWorker Evolves into a Unified Platform for Traditional and Next-Gen Backup and Recover*, now at <http://www.clipper.com/research/TCG2010014.pdf>.

² See [The Clipper Group Navigator](#) dated June 8, 2010, entitled *EMC NetWorker and Avamar – An Integrated Pair for Traditional and Deduplication Backup* and available at <http://www.clipper.com/research/TCG2010028.pdf>

to 30 times or more, which makes disk economical as a backup storage media.

Avamar deduplication identifies and eliminates redundant data at the client (typically a server) before it is sent over the network to a backup server or storage node. After an initial full backup, all subsequent backups transmit only changed, sub-file data. Avamar can identify new and changed data, not just within a single client, but also across an entire data protection domain. Client deduplication generally is much faster than traditional backup because less data is backed up each time. This reduces network traffic and storage capacity requirements. Even with the additional deduplication processing at the client, backup jobs are faster as long as the server's incremental rate of data change is not too high (i.e., less than 10 to 20% daily). Client deduplication is particularly effective for remote and branch office backup consolidation, server virtualization, desktops and laptops, and file servers.

Data Domain deduplication storage systems identify and eliminate redundant data during backup as it is ingested into the system. DD Boost distributes parts of the deduplication process to the NetWorker storage node. Deduplication storage systems are virtually plug-and-play solutions for existing backup processes and policies. They do not require a radical change to the backup infrastructure. It is also effective for transactional applications with a higher rate of data change and high-priority recovery for host servers with large volumes of data that require the fastest recoveries.

Deduplication software and storage systems are complementary technologies, and enterprises may choose to deploy one or both since they possess different advantages.

Conclusion

DD Boost for NetWorker is a must-have option for anyone with Data Domain and NetWorker in their backup infrastructure. It is also a significant advantage for any customers who may be considering an EMC solution featuring both of these products. The enhanced performance, replication management, and simplified administration make for a very powerful feature combination. Through clever engineering, DD Boost actually reduces the impact on the backup server and LAN. The management integration means you can run it all centrally from the NetWorker Management Console,

scheduling deduplication backups and even replication between Data Domain systems.

In the broader context of integration with EMC Avamar and EMC Data Domain, EMC NetWorker has positioned itself as a hinge swinging from traditional to next-generation deduplication backup techniques. For enterprises interested in making this transition, the combination of NetWorker plus Avamar and/or Data Domain will ease the journey.



About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

Michael Fisch is a Senior Contributing Analyst for The Clipper Group. He brings over 14 years of experience in the computer industry working in marketing, sales, and engineering, the last nine of which he has been an analyst with Clipper. Before Clipper, Mr. Fisch worked at EMC Corporation as a marketing program manager focused on service providers and as a competitive market analyst. Before that, he worked in international channel development, manufacturing, and technical support at Extended Systems (since acquired by Sybase). Mr. Fisch earned an MBA from Babson College and a Bachelor's degree in electrical engineering from the University of Idaho.

- ***Reach Michael Fisch via e-mail at mike.fisch@clipper.com at 781-235-0085 Ext. 211. (Please dial "211" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, The Clipper Group Explorer, The Clipper Group Observer, The Clipper Group Captain's Log, The Clipper Group Voyager, Clipper Notes, and "*clipper.com*" are trademarks of The Clipper Group, Inc., and the clipper ship drawings, "*Navigating Information Technology Horizons*", and "*teraproductivity*" are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

After publication of a bulletin on *clipper.com*, The Clipper Group offers all vendors and users the opportunity to license its publications for a fee, since linking to Clipper's web pages, posting of Clipper documents on other's websites, and printing of hard-copy reprints is not allowed without payment of related fee(s). Less than half of our publications are licensed in this way. In addition, analysts regularly receive briefings from many vendors. Occasionally, Clipper analysts' travel and/or lodging expenses and/or conference fees have been subsidized by a vendor, in order to participate in briefings. The Clipper Group does not charge any professional fees to participate in these information-gathering events. In addition, some vendors sometime provide binders, USB drives containing presentations, and other conference-related paraphernalia to Clipper's analysts.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.