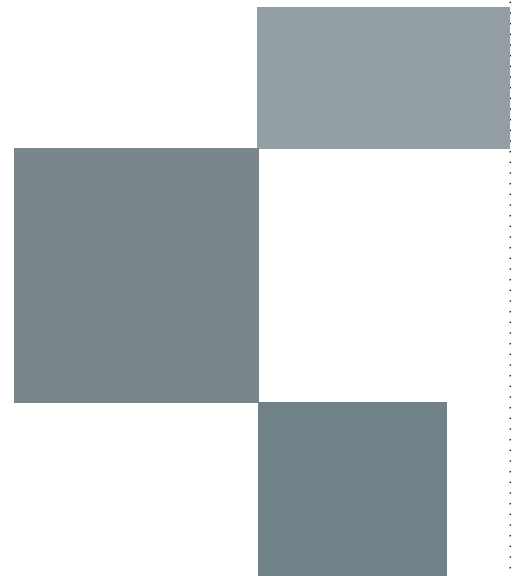




Capacity
Optimized
Storage

Data Domain DD400 Enterprise Series

Capacity Optimized Storage Appliances



Data Domain DD400 Enterprise Series

Data Domain™ OS (DD OS) 3.0, is the software heart of the Data Domain DD400 Enterprise Series. Incorporating Capacity Optimized Storage (COS) technology, DD OS provides the industry's fastest, best-protected backup/recovery storage system. Under control of leading backup software such as VERITAS NetBackup™, CommVault Galaxy™ and EMC/Legato Networker™, DD OS averages a 20x compression effect over time, enabling the industry's most cost-efficient disk-based backup and recovery storage.

Use model: The DD400 Enterprise Series restorers replace or supplement onsite tape library systems as backup-to-disk targets with higher speed and more reliable recovery and, by adding Data Domain Replicator software, provide network vaulting for disaster protection using a tiny fraction of the bandwidth needed by any conventional network replication method.

Benefits

- **Economy:** Order-of-magnitude fewer disks than conventional disk storage through the DD OS 20x Global Compression.
- **Simplicity:** Appliance model, industry standard interfaces (NFS/CIFS) supports all leading brands of enterprise backup software.
- **Fast Restores:** File restores are orders of magnitude faster than tape libraries and they can be done safely by end users in most cases with no administrator involvement.
- **High efficiency network vaulting:** Highly efficient for Wide Area Network (WAN) replication to disaster recovery sites.
- **Data integrity:** A higher level of data protection than conventional file systems or tape systems through the DD OS verifiable recoverability and data self-healing features.





INTRODUCTION 3

THE DATA DOMAIN DD400 ENTERPRISE SERIES.. 4

REQUIREMENTS FOR BACKUP STORAGE 5

TAPE: ADVANTAGES AND CHALLENGES..... 5

DISK: ADVANTAGES AND CHALLENGES 6

SPECIAL REQUIREMENTS FOR BACKUP AND
RECOVERY 7

DATA DOMAIN DD400 SERIES 8

OVERVIEW 8

HIGH PERFORMANCE FOR BACKUPS, RESTORES
AND TAPE COPIES..... 9

CAPACITY OPTIMIZATION FOR COST-EFFECTIVE
RECOVERY STORAGE 11

VERIFIABLE RECOVERABILITY & SELF-HEALING
DATA 15

OFFSITE DATA STORAGE AND DISASTER
RECOVERY 20

SIMPLE INTEGRATION WITH AN EXISTING BACKUP
ENVIRONMENT 23

SUMMARY 25

TECHNICAL SPECIFICATIONS 26

REFERENCES 27

ABOUT DATA DOMAIN 27

Introduction

Data is a key corporate asset and the volume of data continues to grow unabated, by some estimates as much as 100% per year or more. Increased data availability and protection are basic requirements for today's IT organization. A reliable and efficient data protection architecture that can meet ever-shrinking backup and recovery windows and incorporate offsite storage is a fundamental requirement.

Although using disk technology for backup storage allows faster access and better reliability than tape, tape storage has historically been far more cost effective than disk storage. For disaster recovery, most companies have been forced to use a tape backup solution for both economical and operational reasons. Disk storage technology is not designed for, and rarely features a removable media option. However, the notion of data traveling off site on trucks has come under intense scrutiny after widely publicized losses of media also resulted in data privacy concerns.

ATA-based disk arrays are commonly used as staging platforms to briefly hold data as a cache before it is written to tape. But ATA RAID is still about 5x-the cost of equivalent tape storage. So although ATA RAID as a short-term staging platform takes advantage of disk for backup performance, ATA RAID does not improve recovery performance or disaster recovery as restores still rely on tape storage. Most data has to move off expensive disk quickly to keep it affordable.

Staging strategies will provide temporary relief, but do not fundamentally change the data protection architecture which by necessity is still tape based with all its drawbacks and vulnerabilities. A disk based solution must store many versions of backup copies, enable restore operations to leverage disk's random access capability, and allow automatic network vaulting for offsite storage. All the features must be delivered with an economy similar to that of a tape-based architecture. Most importantly, the solution must ensure improved reliability and data integrity.

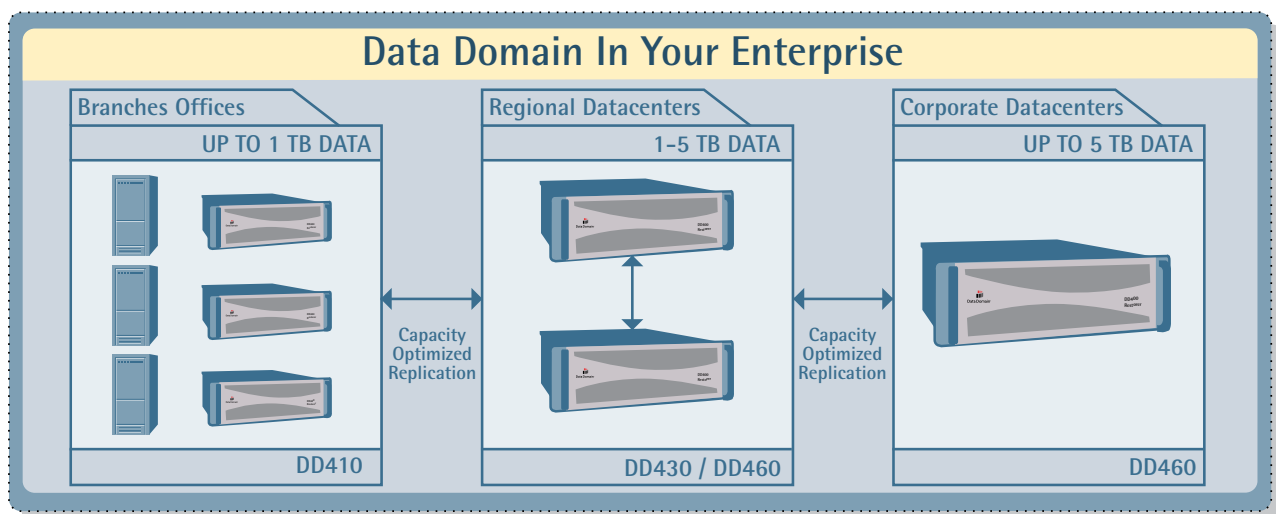


The Data Domain DD400 Enterprise Series

The Data Domain DD400 Enterprise Series of Restorers was designed from the ground up to store backup data and enable automated offsite replication, all driven by standard enterprise backup software. As high performance, online backup appliances, DD400 Restorers enable faster backups, restores, and replication to meet shrinking backup windows and rapid recovery requirements. Restorers integrate seamlessly with your existing backup/recovery processes, architectures, and replication to offsite storage, fully leveraging investments in backup software and training.

DD400 Restorers simply plug into and improve the existing backup infrastructure at the same point in the backup network topology as a staging or caching backup-to-disk target. Define a restorer to standard backup software as a file or disk type device. With a modest change to the backup software policy definitions, use a restorer to backup and recover locally and to replicate data as required for offsite copy management, all under the control of standard backup software.

What makes a DD400 Restorer different? It's the Data Domain DD OS software. DD OS offers unprecedented levels of protection, verifiability and self-healing capabilities unavailable in conventional disk or tape systems. And with its unique Capacity Optimization technology, the DD OS can store several months of recovery copies in an extremely small number of disk drives, lowering the price/GB and greatly simplifying administration.



Requirements for Backup Storage

Until recently, high cost made disks economically unsuited for backup and recovery storage. Disk-based data protection methods such as replication were once reserved exclusively for critical data with the highest availability and recovery requirements, but even that data has historically been backed up to tape. Moreover, to provide a historical repository and to recover from user or operator errors, such as accidental deletion of files, multiple versions of data must be backed up over time, magnifying the cost problem.

With falling disk prices and cost-effective SATA-based disk arrays, disk-based storage and network replication add a new tool for protecting data at an affordable price point. However, simply inserting cheap disk into the process does not provide better data protection or optimize the total cost of ownership (TCO). The next section examines special requirements for an optimized disk-based solution.

Tape: Advantages and Challenges

Tape has been the de facto medium for storing backup data and for transporting replicated data offsite for disaster protection. With increasing capacity and performance, tape allows storing multiple copies or

versions at a low cost.

Unfortunately, tape also has some drawbacks – it is optimized for backup, but not for restore.

With the increasing transfer rates of tape drives, backups need to be carefully staged to stream the tape drive to avoid the “shoe shine” effect of starting, stopping, and repositioning the tape. Incremental backups only worsen the problem since incremental backups do not generate enough data. Thus, streaming a tape drive typically requires “multiplexing,” the blending of concurrent backup streams from multiple clients.

By maximizing tape drive utilization, multiplexing helps backup performance, but slows down restore performance. The extra time required for restores is to read backup images and skip data belonging to other backup clients.

However, the biggest challenge for tape is unknown data integrity. A backup process may have completed successfully, but verifying the data on all of the tapes is near impossible without doing an actual restore. One bad tape can cause a restore operation to fail and render the entire series of tape media useless, and often is not discovered until an actual restore operation.

Disk: Advantages and Challenges

Disk storage offers several advantages over tape. First, unlike tape drives, disk arrays do not need a steady stream of data. Even incremental backups that generate small amounts of data do not create a “shoe shine” effect.

Second, disk arrays can simplify and speed up the overall backup process by allowing the administrator to perform less frequent full backups without suffering a performance penalty or increasing restore risk. Despite shrinking backup windows, when using tape, frequent full backups are performed to minimize the number of tapes required for restores, simplifying the recovery process. The number of tapes required for a restore increases with incremental backups, which increases the time required for restore and the risk that one of the tapes is unrecoverable. Disk allows the administrator to shorten backup windows.

Disk also makes off-site recovery copies easier and more efficient. Instead of multiplexing data from multiple backup clients on one tape, disk-to-tape vaulting allows a straightforward organization of recovery copies for each client, thus speeding up the recovery process. In addition, disk-to-tape copying provides more flexibility than tape-to-tape. When making tape-to-tape copies, both primary and clone tape drives are unavailable for other backup or restore operations. In contrast, disk allows

simultaneous access, accepting backup and restore operations while copying data to tape.

Most importantly, disk is superior for recovery – in reliability and performance. Disk-specific technologies like RAID make disk a more reliable medium than tape. As mentioned, one bad tape in a sequence can cause an entire restore operation to fail. With RAID protection, a restore can continue and complete successfully even with a failed disk. (While vastly better than tape, RAID is not by itself sufficient to guard against all issues, as explained below.)

Moreover, according to Strategic Research, 87% of all restores are single file recoveries, not full system recoveries. As a random access device, disk enables much faster single file recoveries. Average access times for disk are measured in milliseconds. In contrast, average file access times for tape, a serial access device, range from 27 to 73 seconds. If a restore requires tape location, loading, and unloading operations, the overall “time-to-data” is even worse. It’s no contest.

Despite the advantages, the biggest limitation for standard disk storage for backup retention continues to be cost. Even with dramatically falling disk prices, tape has been the most economic choice. That is why it is still around. The cost equation is simple. Imagine storing four weeks of weekly full and daily incremental backups. Assuming incremental backup sizes are approximately 5% of the original data set,

the storage requirement is five times the original data size. Using primary storage subsystems such as EMC Symmetrix for backup storage is obviously cost prohibitive. At \$10-\$20 per GB, even SATA RAID arrays are still far more expensive than \$1-\$5 per GB for tape-based backups.

Ensuring continued data integrity is also still a challenge for conventional disk-based storage. After all, storage systems can fail or lose and corrupt data. That is why we backup data in the first place.

Special Requirements for Backup and Recovery

A backup storage solution should combine the economy of tape with the usability and speed of disk. At the same time, it must overcome the disadvantages of both tape and conventional disk storage arrays.

Special requirements for backup storage are:

- **Economy:** The solution should have a cost comparable to tape automation. We don't want to spend more money; we just want more bang-for-the-buck that we are already spending. It must also allow for weeks and months of retention so all restores can be done from disk.
- **High Performance:** Shrinking backup windows and growing volumes of data demand high

backup performance. At the same time, increased importance of data to day-to-day operations requires high restore performance. The performance of a backup storage solution must be optimized for both operations. Tape technology has kept up with backup performance needs but at the cost of unacceptably slow restore performance.

- **Data Integrity:** The purpose of backup is to recover data. How do we know the backup we just made is any good? Even after a backup job is successfully completed, it is critical to verify that the resultant backup image is recoverable. It cannot be taken for granted, given the large number of restore operations that fail outright. Furthermore, to ensure successful recoveries, a backup storage solution must deliver much higher levels of hardware and software protection than conventional RAID and file systems. After all, it is the data store of last resort. The ideal backup storage solution must provide verifiable recoverability and highly resilient storage.
- **Low Disruption:** Few can afford to rip out existing systems and start over. What's required is a simple-to-use and easy-to-integrate solution into existing, standard backup/recovery operations environment: According to Gartner, backup and recovery account for the largest portion of storage total cost of ownership (TCO) at 30%. Today more than ever, IT

organizations are asked to “do more with less.” An optimized backup storage solution must simplify the backup and recovery process and enable IT organizations to leverage existing investments, using standard interfaces.

- **Automatic offsite replication:** Disaster recovery must be part ‘n parcel of the data protection architecture. It can’t be an afterthought. Backups should be vaulted over a Wide Area Network, or we will be forced to make tapes again and transport our corporate data on trucks. Highly publicized data losses and data privacy violations in recent times speak volumes in this regard. An efficient WAN replication strategy uses minimal bandwidth to migrate backups over the network while simultaneously performing other operations such as backups, restores or clones.

Data Domain DD400 Series

Overview

Data Domain DD400 Restorers are disk-based backup storage appliances. While built on serial ATA disk technology, DD400s are not just another cheap disk array with RAID. The DD OS, with its revolutionary Capacity Optimized Storage technology and Data Invulnerability Architecture, provides unprecedented data protection with a resulting cost-per-GB that is dramatically *lower* than cheap disk, approaching the cost of tape automation, and in some cases approaching the cost of tape cartridges. DD400s are designed to meet the unique demands of backup and recovery storage.

- **High performance for both backup and restore:** Throughput with a single DD400 controller is up to 290 GB/hour, similar to current LTO-3 tape drives, and many primary storage systems. However, unlike a tape drive, a DD400 does not require constant streams of data for the best backup performance. Multiple backup streams at varying speeds can be sent to one DD400. A DD400 also enables fast single file restores by taking advantage of the random access nature of disk. By configuring multiple DD400s, it is possible to scale cumulative throughput to any required level.
- **Economy:** The DD OS dramatically reduces the storage required for backup data by pooling redundancies within backup images

and storing only unique data patterns. This allows the DD400 to not only detect and eliminate storage of duplicate files but also to detect and eliminate repeated patterns within and across files. With its unique Global Compression technology, the DD OS delivers an effective compression ratio of 20:1 over time. As a result, a DD400 is an order of magnitude smaller, simpler and easier to administer than cheap disk arrays used for backup storage.

- **Data Invulnerability:** A DD400 is designed to prevent, detect, and heal from hardware or software failures to ensure data integrity and restorability. The DD OS file system and RAID design were built from the ground up to offer fault protection, detection and correction from software flaws and disk errors in a much more rigorous way than general purpose disk storage or file systems.
- **Easy to use and integrate into an existing backup software environment:** To a UNIX, Linux, or Windows storage administrator, a DD400 is a familiar sight with its industry standard NFS and/or CIFS interface. A DD400 fits easily into an existing backup software environment and supports leading enterprise backup software from vendors such as VERITAS, EMC/Legato and CommVault. The DD OS even enables some features of backup products that are often overlooked, such as end-user-initiated file recoveries.

High Performance for Backups, Restores and Tape Copies

The DD400 Series was designed with three performance requirements in mind:

- Faster backups to meet ever-shrinking backup windows
- Faster restores to meet ever-shrinking tolerance for downtime
- Faster and easier replication to off-site storage locations

Backup Performance

A DD400's performance is comparable to high-end serial ATA RAID system in backup/restore target applications. Unlike a tape drive, a DD400 does not require backup servers to send constant-speed streams of data for best performance. With tape drives, multiplexing is used to compensate for slow clients, multiple slow networks, and short backups such as incremental backups. With the DD400, multiple, concurrent backup streams at varying speeds can be sent to one DD400. For incremental throughput and capacity, multiple DD400 units will scale to transfer any quantity of data in the available backup window.

Recovery Performance

Recovery performance is where a DD400's advantages stand out most. Assume a typical backup schedule of a full backup on Saturdays followed by daily incremental backups. With tape-based backups, a full restore on Friday morning requires reading all of the tapes serially, and loading and unloading the tapes. If the required tapes were not available in the tape library, the restore operation would first

need to locate the tapes, which might require hours, if not days. Multiplexing would further degrade performance because of the need to read tapes and skip data belonging to other backup clients. On the other hand, a DD400 can sustain high performance when responding to backup software's requests for appropriate restore images no matter how scattered the images.

Consider a similar but more common case – a partial restore on Friday morning where only a few files need to be recovered. With tape-based backups, even a partial restore may require several tapes. Again, the restore operation would take longer if any of the needed tapes were not in the tape library. Given the average file access and media loading times, time to begin reading the needed files is measured in *minutes*. With a DD400, such time to begin reading is measured in *milliseconds*. By taking advantage of the random access aspect of disk, a DD400 enables faster restore performance than tape.

DD400 Online Retention vs. Staging

Because a DD400 can efficiently store many months of backup data, most, if not all restores are performed from the DD400. That is not the case with a disk staging device, where most restores will still be performed from tape media. In some backup systems, disk is used as a cache for tape. In a cache scheme, a backup copy is first directed or staged to disk. After the backup image is copied to tape, the image is deleted to make room for the next backup. The process leverages the advantages of disk to improve backup performance. However, restore operations

still require tape, incurring performance penalties of locating, loading, and serially accessing the right set of tapes.

DD400-to-Offsite Replication for Disaster Recovery

The DD400 makes off-site replication or network vaulting much easier, faster and affordable. Optional Data Domain Replicator software takes advantage of the unique DD OS features to create a simple and efficient mechanism that replicates backup data asynchronously across a network between two DD400 appliances. With Replicator, data is backed up to local DD400 using standard backup software. The local DD400 then acts as an "originator" and replicates the data over a TCP/IP LAN or WAN to a remote DD400 replica.

Thanks to Capacity Optimization, there is a massively reduced amount of data sent over a WAN by 95%+ compared to replication from any other disk-based backup system. If necessary, data can be restored across the network from the replica with the same efficiency or data from the replica can be restored at the remote site. A replica DD400 at a remote site can also vault data to tape at the remote site for long-term archiving. Because of this, the effective replication performance of a pair of Restorers is much higher than the available bandwidth of the link. ESG Labs has shown effective performance in this case of more than 450 MB/sec., since only the unique data segments need to be exchanged for the entire image to be transferred.

DD400-to-Tape Copying

When copying data to tape for long-term archiving, a single DD400 is fast enough to stream a high-end tape drive for best performance. In addition, a DD400 is available and accessible for backup or restore operations while copying data to tape. In contrast, with tape, if a file must be restored from a DD400 while making an archive tape copy, the copy process does not need to be interrupted.

Combining its high throughput, capacity to store many weeks of backup data, random access capability, and high-speed replication over a network, a DD400 improves performance of the entire backup and recovery process.

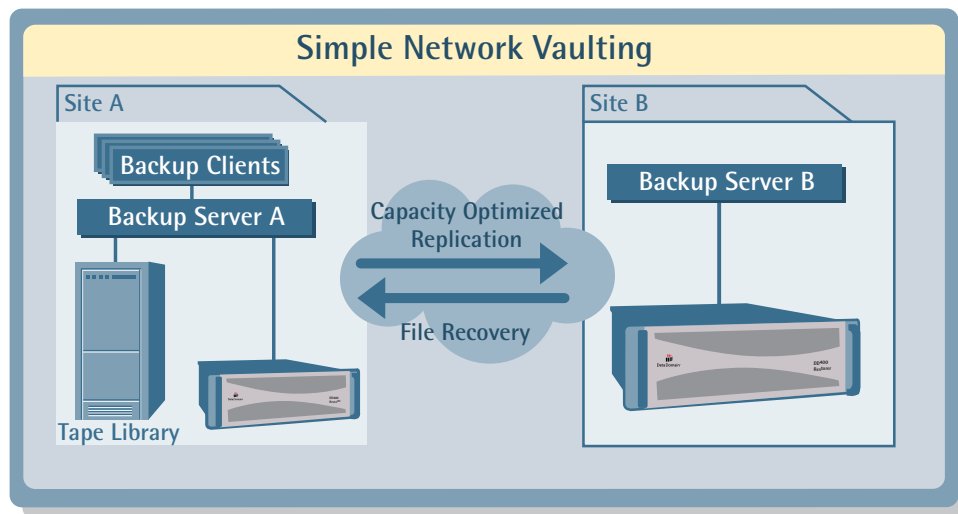
Capacity Optimization for Cost-Effective Recovery Storage

Backups contain a lot of redundant data, especially when comparing one weekly full backup image with

subsequent full backup images. Although incremental backups capture only changed files, incremental backups generally also contain mostly redundant data blocks.

The DD OS Capacity Optimization approach pools redundancies within backup images and stores only unique data segments. When data is written to a DD400, the data is broken into variable-length segments, or sequences of bytes. The DD OS real-time compares each segment to all segments already stored. The process ensures that each unique segment is saved only once. Thus, the DD OS is able to detect duplicate files, redundant patterns within and across files, and even repeated patterns within blocks. The stored data is smaller by more than an order of magnitude than the amount backed up.

A key to the effectiveness of Capacity Optimization is independence from data format. DD OS's implementation, Global Compression, is based on



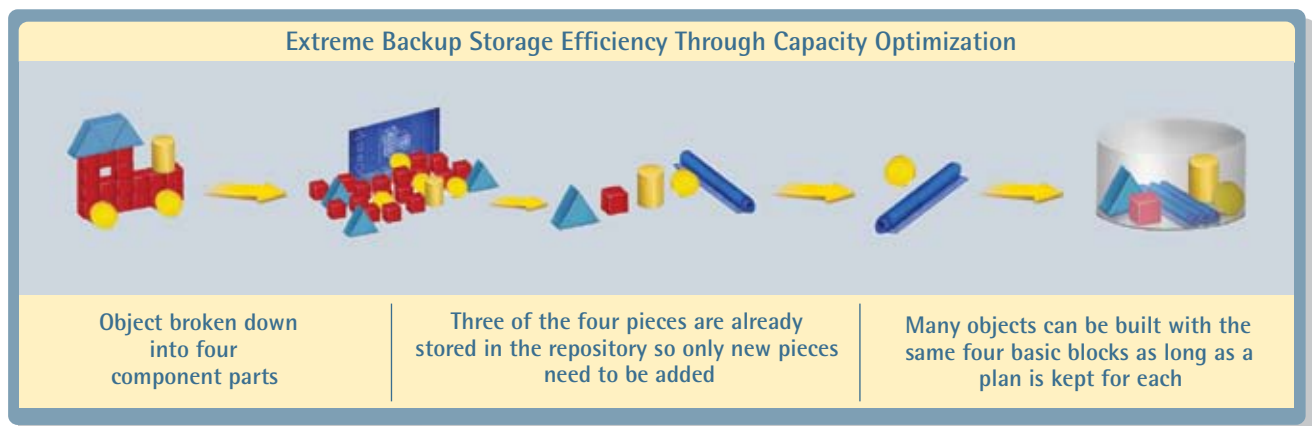
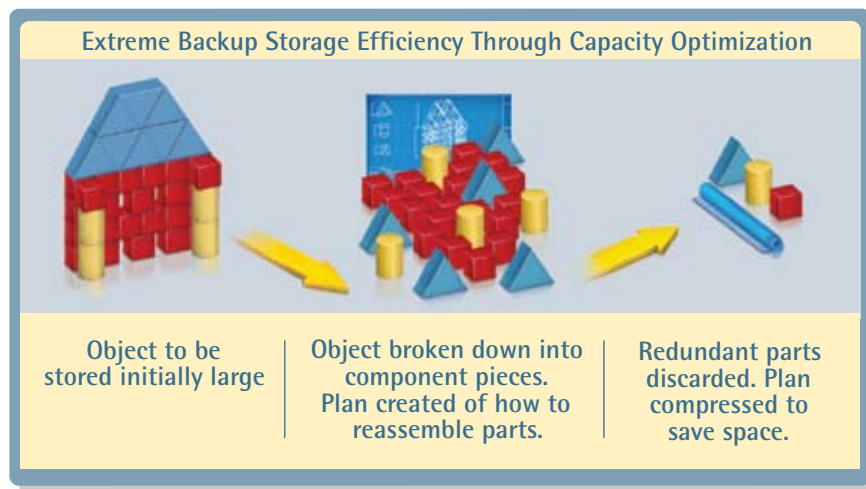
All backups are targeted to disk device at Site A;
 All backup data is replicated asynchronously to Site B;
 File recovery can take place from either Site A or Site B,
 and tape is relegated to long-term archival.

analysis of data contents and repeated patterns. Global Compression can be applied to any type of iterating data – structured data such as databases or unstructured data such as text files, data stored in file systems or raw volumes. With the Global Compression algorithm, the DD OS is highly efficient for backup copies, regardless of data format.

While many of the central ideas underpinning

Capacity Optimized Storage are not new, and have been published in academic circles for decades (see list of references later in this document,) Data Domain’s implementation and intellectual property is in the outstanding performance of the optimization algorithms, allowing the benefits of massive data reduction to accrue virtually transparently to the backup/restore process.

Data Domain’s Capacity Optimization: How it Works



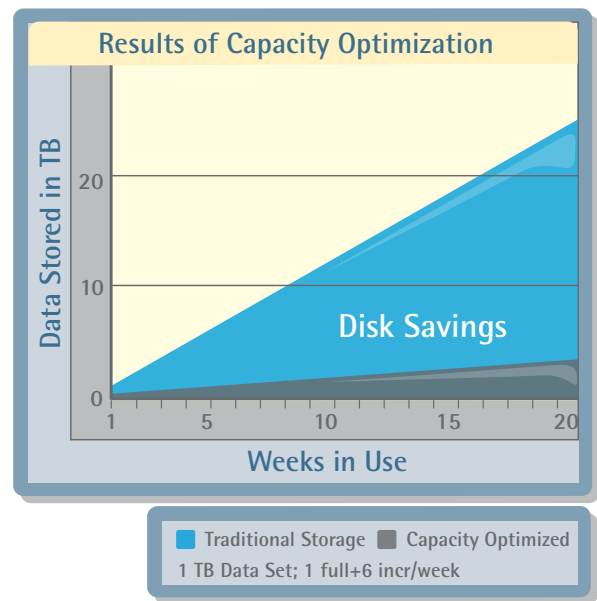
The Global Compression algorithm analyzes all data stored within a DD400 as it is stored and is independent of data formats and alignments. Optimization benefits increase as more and more data is stored over time. The single biggest lever on compression effect is backup policy. Full backups have much more redundancy than incremental backups; more fulls can increase the compression effect dramatically. But even with incremental backups, the Global Compression approach can offer 80% - 90% data size reduction.

Over a four-week period, 10:1 data compression is expected assuming weekly full and daily file-level incremental backups. Over an eight-week period, 14x compression is expected. A 20x compression benefit is to be expected after about 18 weeks. With a full-daily backup policy, the best-practice approach recommended in Oracle and Exchange backup/recovery, much greater compression effects happen faster, often exceeding 10x after the first week and 20x after the second.

Global Compression has been applied to various data types such as Oracle and other SQL databases, SAP data, text files, home directories, engineering data such as software development and semiconductor design, UNIX binaries, PC applications, Microsoft™ Exchange, and Microsoft Office documents.

Also affecting the compression factor is the data itself. A backup image that consists of many duplicate files or similar files (where a file is copied several times with minor changes) benefits greatly from Capacity Optimization. Some kinds of data

do not iterate, and do not benefit as much, such as seismic mapping data or satellite telemetry data that includes a static binary image. If data has no redundancy to pool, the DD OS approach still reduces the data footprint more than most other general purpose compression or duplicate-file-elimination algorithms.



The first full backup results in 3-4x data reduction; File level incrementals will see 6-7x and subsequent full backups will see 50-60x data reduction ratios. The aggregate with Weekly Fulls and Daily Incrementals is 20x.

Global Compression vs. Snapshots or Block-Level Incremental Backups

Global Compression enables highly efficient storage of backup data. How does it compare to other storage-efficient technologies such as file system based snapshots or block-level incremental backups?

A file system snapshot is based on copy-on-write techniques where a snapshot consists of changed file

system blocks. Similar to snapshots, block-level incremental backups capture only changed blocks (either at the file system level or at the application level, such as databases). By storing only changed blocks, the technologies are more storage efficient than other techniques such as third mirror break-off or file-level incremental backups. However, note these major differences between Capacity Optimization and other technologies.

- **Heterogeneity:** Capacity Optimization is data format and storage independent, applies to structured or unstructured data, and supports any file system or raw volumes. In contrast, snapshots are file system- or volume manager-specific. Only data within those systems benefit. Block-level incremental backup techniques are also specific to underlying mechanisms that keep track of block changes in the file systems or applications such as Oracle™ databases.
- **Storage efficiency based on backup software integration:** Other technologies are integrated with backup software. However, backup software takes advantage of them differently, affecting storage efficiency. Most backup software products leverage snapshots' capability of providing stable, point-in-time views of file systems, not the ability to capture changed blocks. When backing up from a snapshot, backup software saves entire files, losing storage efficiency. Similarly, block-level incremental backups, though much more efficient than file-level incremental backups, still require full

backups which contain mostly unchanged data blocks.

In contrast, Capacity Optimization is applied to data from backup servers as the data is being stored on a DD400. Whether backups are taken from snapshots or contain only block-level changes, the DD OS can analyze repeated patterns, resulting in even higher levels of efficiency.

- **Storage efficiency based on data content vs. location:** Capacity Optimization provides much more efficient and effective storage for backup data than snapshots or block-level incremental backups because of data content analysis and a finer granularity of changes. The DD OS tracks changes or finds new segments based on content and compares against all data retained within a DD400. On the other hand, file system snapshots or block-level incremental backups track changed blocks on a per file basis. Consider the following simplified examples:
 1. **Duplicate files:** Suppose file1, is copied as file2. After backing up only unique segments in file1, a DD400 does not save any additional segments for file2 since the content of file2 is identical to that of file1. From a snapshot or block-level incremental backup point of view, file2 is a new file and all of its data blocks are new, thus requiring storage for all of the blocks in both file1 and file2.



2. **An identical change in multiple files:** Suppose both file1 and file2 are updated by appending the same data. Global Compression examines the new data. If the new data is unique, it is saved once, although both files are updated. Unlike Global Compression, the snapshot and block-level incremental backups track and store the change per file.
3. **Multiple, identical updates within a file:** Assume file2 is modified with the same data multiple times. As Capacity Optimization can detect repeated patterns within file2, only unique portions of the change are saved once. Similar to the previous example, with snapshots or block-level incremental backups, each change is stored separately.
4. **Change in alignment in a file:** After insertions or deletions in a file, the data is no longer aligned with the original layout of the file. As the Capacity Optimization algorithm is based on content, not location, it recognizes the unchanged data. However, to the snapshot and block-level incremental backup, every block after the insertion or deletion appears to be new, requiring storage for unchanged blocks.

For recovery copies, Capacity Optimization requires less storage than snapshots or block-level incremental backups. Since a DD400 supports standard backup software, snapshots and block-level incremental backups can be sent to a DD400 with Global

Compression to gain additional storage efficiency.

Verifiable Recoverability & Self-Healing Data

The DD OS offers significantly stronger file system and data integrity verification and capacity to self-heal than even high end enterprise primary storage systems. After all, you are backing up your primary storage systems for a reason – they have vulnerabilities. The backup storage system should perform at a higher level of protection than primary storage.

- In a primary storage file system, logical consistency checking is always a risk zone. If a software bug causes wrong data to be written, a block pointer, bitmap, or link count can go bad. Usually, the best way to reveal such problems is to run a file system check (such as `fsck`) after un-mounting a file system. If the file system is storing backup data, the fault would otherwise not be noticed until the next time the data is read: during a recovery, which is not an acceptable time to find out.
- In high end enterprise RAID systems, data blocks are reasonably safe from disk errors given enough time. Blocks that have errors on disk can be corrected by a scrubbing operation that compares the data against checksums when data is read. In backup storage that might be weeks later. If a disk failure happens in the meantime, data could be lost. Many

RAID systems do not perform soft-error corrections at all. And when corrections are performed, they correct only data segments, not file system flaws. And in the end, these checksums can fail; if an error creates the same checksum that the correct block would create, it would go unnoticed; checksums are a kind of hash, and this is a “hash collision.” Verification is only as strong as the weakest checksum in the system.

Backup data is most valuable within days of the backup. Backup data is also not read very often, but when they are, time matters – the need for data recovery means a human or system failure somewhere else has already occurred. Waiting until the next boot time or taking the system offline to review file system consistency in a recovery store adds unnecessary risk. Depending on weekly scrubbing operations for data integrity checks means finding bad news when you can least afford it.

The DD OS includes:

- File system consistency verification & self-healing: initial, ongoing, and online
- File system software fault protection
- Data integrity verification & self-healing: initial, ongoing, and online
- Sound data protection fundamentals, such as:
 - Data fault protection from disk errors and power failure inconsistencies
 - Integration with existing processes to complement tape for offsite support.

With the above features, a DD400 is much more likely to receive incorrect data (from primary storage or storage fabric errors) than to recover the data incorrectly. Periodic recoveries are still prudent to test the data from the point of view of the backup software and end user applications.

File System Consistency Verification & Self-Healing: Initial, Ongoing, Online

The weakest link in primary storage relative to rapid integrity verification is the file system. Traditional enterprise file systems, while extremely high quality, are still software and subject to bugs. A bug can cause a pointer, bitmap, or link count to be incorrect. Such problems often do not appear until (1) corruption brings the system to a halt, or (2) an offline file system check reveals the error.

In the DD OS, the consistency of new file system metadata and integrity of the data itself is verified after the backup completes, initially within hours. All data is then *continuously* rechecked online in the background.

DD OS Initial End-to-End Verification Process

1. Strong Checksum on New Data
2. Write to Disk
3. Read back from Disk through the File system
4. Compare to Checksum

Initial end-to-end verification. Strong checksums are calculated on receipt of data by the DD OS, after which data is stored to a DD400’s battery-backed NVRAM and then to disk. Within hours, the DD OS regenerates the checksum and *reads the data back*

through the file system from disk to make sure data matches what was stored, validating data integrity and logical consistency. The DD OS can warn the operator about a backup error in time to address it. Conventional file systems do not check consistency online at all.

Ongoing checks and reporting. In the DD OS file system, metadata records and in the content data itself, the DD OS file system is also self-describing. Metadata includes reference information about where the data is supposed to be with respect to the addresses of other data in the system. The reference information enables active, ongoing file system consistency checking.

Most problems can self-heal; all can report quickly.

A recent, separately stored snapshot copy of the file system is maintained internally; in many cases, correct metadata in that copy can be used to upgrade flaws in current metadata. For example, the index that maps unique data segments to virtual files includes metadata with data on disk, so data can completely self-heal from the on-disk information. If a higher level consistency problem exists that requires un-mounting the file system to repair, thanks to early and active verification, the DD OS provides notification while you have time to plan for repairs.

Append-only File System for Software Fault Protection

The DD OS file system is much simpler than those of most enterprise file systems, minimizing the potential for error.

Most file systems are optimized for random block updates and low latency. Data structures that facilitate such features are somewhat complex and subject to inconsistencies. Correctly managing the data structures (block address pointers, bitmaps, and link counts) is hard to do well and especially difficult across system faults. Due to the complex software, bugs are hard to find.

The DD OS was built with a very different design objective. Backup files tend to be large, input as whole files, and are typically not modified by backup software. Backup files are written, read, or deleted, and systems have daily idle time. The DD OS is developed to the requirements of the backup process. As a result, the DD OS is able to deploy a storage layer that is significantly simpler and more faulttolerant than general purpose storage systems.

The DD OS writes data in a sequential log structure. Once a block is written, it cannot be updated or overwritten until a formal clean command completes (safely moving in-use blocks to the end of the log, then only freeing blocks no longer used by files after the blocks are deleted by backup software). The process means that the DD OS can never become confused and overwrite valid data; valid data is safe once it is on disk.

By eliminating complex data structures, the DD OS can also offer much more rigorous and online file system and data integrity verification and reduce the overall risk of error. If a pointer in a conventional file system says to get the data at block #82, that is what

a request will get. If the pointer should have said #84; a typical file system has no way to check correctness online. The DD OS accesses data more simply and re-verifies data during the read process, so such problems cannot happen.

Data Integrity Verification & Self-Healing: Initial, Ongoing, Online

RAID systems have two approaches to data integrity checking and correction:

- High-end enterprise RAID arrays use checksums at the block segment level, flagging errors when data is read. A few high-end arrays scrub the data in the background well after writing it and correct data using parity images or mirrors. Such strategies enable a verification time lag of days or weeks.

If verification is exercised only when the data is read, errors may be found only during restore. If a disk fault occurs before data is checked in a parity RAID configuration, data may never be recovered.

- Lower end RAID systems do not use checksums or scrubbing at all, so disk segment errors remain as hidden, uncorrectable fault zones.

ATA RAID is more vulnerable to errors than fiber channel RAID because ATA RAID does not assure command queuing in its cache. For example, in a power failure, data acknowledged by the disk as written to magnetic media may not have moved

out of cache, so data could be lost, though the RAID system may believe that the data is written. ATA RAID vendors need to do additional validation to ensure data integrity.

The DD OS offers a timely and complete approach to ensuring data integrity. As described above, the DD OS does an integrity check within hours of a backup, including both data integrity and file system consistency checking. If an error is found at the disk level, the DD OS can correct the soft error by using the DD RAID parity data and remapping the bad block. After the initial recovery data verification, the DD OS also offers continuous background data verification using a very strong checksum. How strong? For example, each of a DD400's ECC-enabled disk drives has a risk of a single-bit error once for every petabyte written (a thousand times or more the capacity of the disk itself); the DD OS's checksum validation is stronger by a factor of more than 10^{20} . Through strong testing, the odds of a data error are much smaller once data is inside a DD400 than the odds of an error before data reaches a DD400. Data Domain still encourages periodic checks of backup images from the point of view of the backup software and the application. As a disk-based system, a DD400 is much more supportive of data testing than tape automation can ever be.

Sound Data Protection Fundamentals

A DD400 also includes a full range of conventional data protection technologies, including:



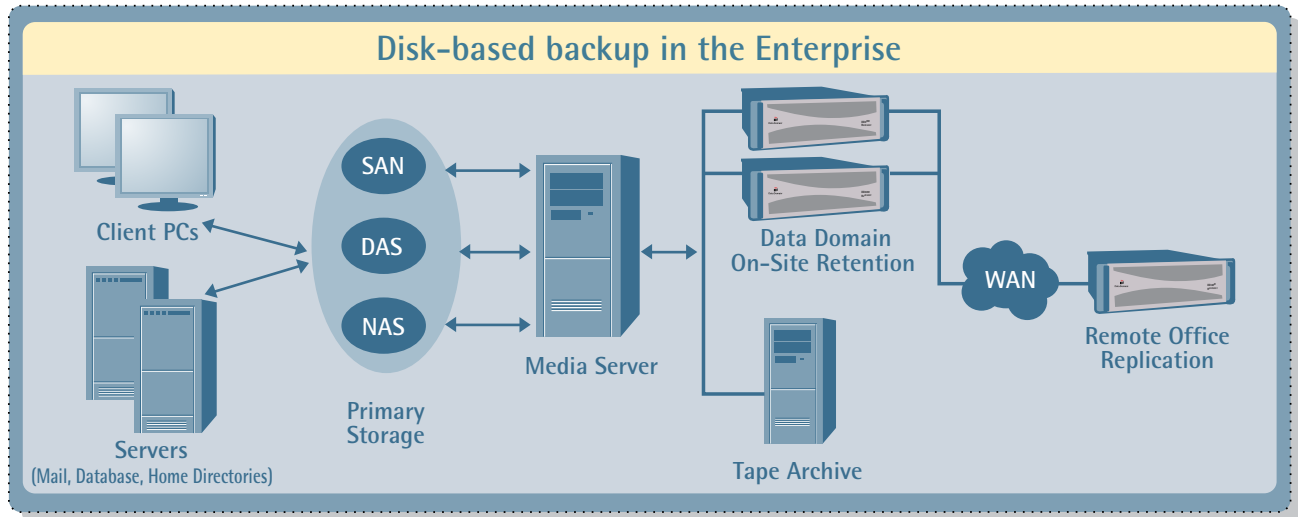
- **Double-Parity DD RAID.** Each of the two DD RAID parity stripes has block checksums to ensure that data is correct. The checksums are constantly used during online verification and when data is read from the DD400. With double parity, the system can fix simultaneous errors on up to two disks. RAID operations are done in the DD OS, not in hardware. If a bad block is found, the system corrects the block using a parity stripe. This is a form of what is sometimes called RAID6. It is significantly safer than mirroring or RAID5 with a hot spare, which is always vulnerable to a second error during reconstruction.
- **NVRAM for storage buffering.** NFS V3 acknowledgements are returned once data is stored by the DD OS in a DD400's battery-backed NVRAM. In case of a power failure, recovery of writes that have not gone to disk requires simply playing back the requests and checking against and overcoming the power-fail weakness of the ATA cache. No metadata updates are committed to disk until all data is already committed, allowing no consistency confusion in case of NVRAM failure. In the event of NVRAM failure, only the most recent write requests are lost; the system is otherwise consistent. To minimize the amount of NVRAM required, NVRAM holds only post-compression data.

- **Data integrity that complements tape support.** Backup software is very accustomed to tape drive failures, understanding very well how to signal operators and initiate restarts. Restore operators are also familiar with such situations.

A DD400 is significantly more reliable than a tape mechanism. In most cases, a DD400 can experience a problem, correct it, and continue a backup uninterrupted. In the rare case of a system fault that is critical, the backup is no worse than with a tape drive fault. Full data integrity is preserved in data received up to that point, but the backup may need to be restarted. Unlike primary storage used as backup storage, a DD400 gives no opportunity for silent data loss – full integrity and consistency awareness allows comprehensive reporting to the operator.

Backup software is expected in most cases to make copies of tapes for offsite storage based on service level requirements. Multiple copies are the ultimate strategy for maximum protection, and DD400s support all leading backup software packages in cloning to tape.

Offsite Data Storage and Disaster Recovery



DD400 Enterprise Series restorers provide support for two types of offsite data protection: capacity optimized replication for full automation driven by standard backup software, aimed at minimizing bandwidth costs, and traditional tape vaulting.

For most companies, tape backup with offsite tape storage is the de facto strategy for disaster recovery. However, with the astronomical growth in stored data, plus the trend toward global, round-the-clock operations, tape solutions by themselves are now inadequate.

In addition, tape technologies are prone to error. Discovering that a critical data tape is unreadable in the midst of recovering from a disaster can cripple a previously successful business. For companies with remote locations, the difficulties of tape may be compounded by lack of remote staff with the

expertise to correctly manage complicated tape backup systems.

Because tape infrastructure, software, and best practices are mature and expensive to change, IT groups concerned about disaster protection are forced to choose between

- Sending tape offsite
- Keeping tape onsite for recoveries
- Making two copies (as if one wasn't bad enough)

Many IT organizations now see network vaulting or online replication as an alternative to offsite tape storage for disaster recovery. With replication solutions, data is copied from local primary disk storage to remote disk storage over a LAN or WAN. Data can be synchronized more frequently for increased protection, and the remote site can be

configured as a complete disaster recovery site where operations can be re-started if the primary site is down for an extended period.

While disk-to-disk replication over a wide area network provides the highest level of disaster recovery capability, it traditionally has some significant cost factors that make it difficult to afford for any but the most critical high-value applications. In particular, the cost of WAN bandwidth can be prohibitive and has not fallen as rapidly as once predicted. In addition to the costs, available software solutions are expensive and often complex to implement and manage. Such factors have inhibited many companies from seeing the benefits of online disaster recovery technology.

See also Data Domain's separate technology note on DD Replicator at:

www.datadomain.com/products/dd-replicator.html

Data Domain Replicator

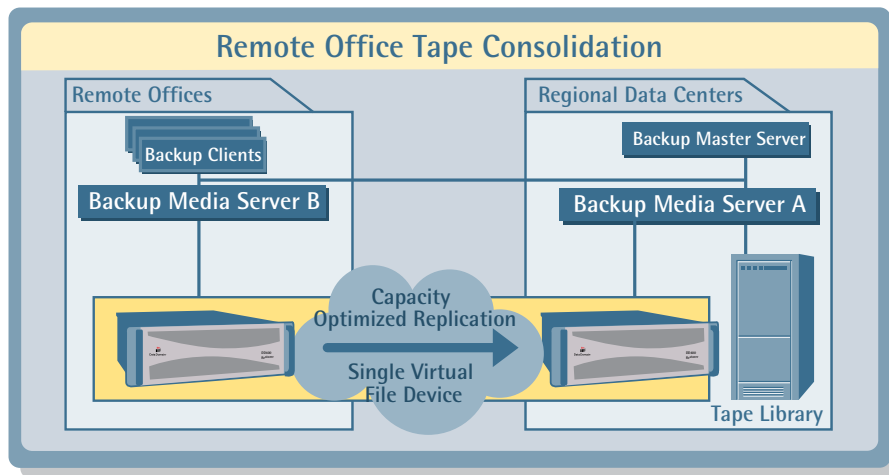
With Data Domain Replicator software, Data Domain substantially decreases the cost of disaster recovery. Data Domain Replicator offers unparalleled data safety with proactive verification and self healing, plus the dramatic reduction in bandwidth utilization derived from 20X Global Compression.

Data Domain Replicator allows you to set up a highly economical online recovery infrastructure. Here is how it works:

Backups are made to a DD400 (*the originator DD400*) at the primary site using standard backup software such as VERITAS NetBackup, Legato Networker, or CommVault Galaxy. The originator DD400 at the primary site is configured to automatically replicate all newly received backup data to another DD400 at a remote site.

During normal operation, the primary site can restore from the originator as necessary and perform other normal operations such as duplicating/cloning backups to tape (for archival, regulatory compliance, or other purposes). Tapes for long-term archiving can also be duplicated from the replica at the remote site. Depending on the nature of the remote site, the complete tape library can be located at the replica site for consolidating necessary tape operations.

The remote site can also be a fully operational data center. If the DD400 originator fails or is affected by a disaster, data can be restored across the network, or a copy of the replica can be shipped back to the primary site for restoration. If the primary site is beyond recovery due to disaster, a backup server can be configured and all data can be recovered at the remote site.



This graphic illustrates backup to a remote office and replication to a regional data center. Restores can be done at any location, and this scenario eliminates tape backup at remote sites.

A compelling attribute of Data Domain Replicator is ease-of-management, especially relative to manual shipment of tapes. Some upfront planning must be done to determine the necessary bandwidth; but otherwise, initial configuration can be done in a few minutes. Once configured, replication proceeds 'lights out' automatically whenever the originator has un-replicated data. The status of replication can be monitored with a few simple commands.

Using Data Domain Replicator for disaster recovery offers the following benefits:

- **Economy.** The high cost of disaster recovery results primarily from storage and network bandwidth costs. Replicator massively reduces storage and bandwidth costs with Data Domain's 20x Global Compression. Data sent over a WAN is reduced by 95%+ compared

to replicating any other disk-based backup system. And backup data travels on a network, instead of on trucks.

- **Highly resilient.** Replicator features continuous verification of data integrity to ensure data can always be recovered. If an error or interruption occurs during transfer, Replicator recovers automatically and resends the data.
- **Simple to set up and manage.** One of the hidden costs of traditional disaster recovery solutions is the complexity of set up and management. Set up Replicator once, and from then on, only exception monitoring is required.
- **Works with standard backup solutions.** Replicator preserves your investment in

software, process and expertise in enterprise-class backup software such as VERITAS NetBackup, CommVault Galaxy and EMC/Legato Networker.

- **Platform independent.** Because Replicator works with your backup solution, it automatically provides data protection for every platform you back up, whether that platform runs Linux, UNIX, Windows, or other operating systems.
- **Data is always accessible.** With some replication technologies, data cannot be accessed while replication is in progress. With Replicator, both the originator and replica remain fully accessible while transfers are in progress.

Simple Integration with an Existing Backup Environment

To a system or storage administrator, a DD400 appears as a Network Attached Storage (NAS) appliance supporting NFS and CIFS over Gigabit Ethernet. But a DD400 is more of a “restorer” than a “filer.” Global Compression allows a DD400 to store 20 times more backup data than a NAS appliance with the same useable disk capacity. Leading backup software packages such as Legato NetWorker and VERITAS NetBackup already support backing up to disk by specifying a directory as a managed device. Simply add a DD400 as a Filesystem or Advanced Filesystem Device in NetWorker, or as a Disk Storage Unit in NetBackup. Update or add a policy to target

that device instead of a tape system. Backups are then targeted to the DD400. NetWorker and NetBackup also manage recovery images on a DD400 by deleting appropriate files when the corresponding recovery images expire.

Once backups are safely stored in a DD400, tape copies are easily made by cloning, duplicating, or vaulting recovery images from the DD400. Again, such features are already supported by backup software. The simple changes needed to use a DD400 require no additional software to be loaded, no special integration with backup software, and no training on a new backup package

Better Use of Backup Software: User-Initiated Restores

As an NFS file server, a DD400 provides additional benefits. One DD400 can be shared across multiple backup servers without requiring additional licenses. But more importantly, one DD400 can be used for simultaneous access, allowing multiple, concurrent backups or simultaneous backup and restore operations.

The flexibility of simultaneous access also allows IT organizations to make better use of existing backup software. Backup software products, such as NetWorker and NetBackup, support user-initiated restores, but this feature is not widely enabled as user-initiated restores can disrupt already scheduled backups by creating tape resource contention. With a DD400, users can restore their own files

without impacting regularly scheduled backups as concurrent backup and restore operations are supported from one DD400. By leveraging the capabilities of both backup software and a DD400, user-initiated restores reduce the administration time needed when recovering from accidental deletion of files.

Support for Backup Software Standardization

A DD400's seamless support for leading backup software allows you to leverage already standardized backup products or to decide on the backup products without being locked into a DD400-specific configuration.

Deploying a DD400: Easy and Seamless

A DD400 requires a simple modification in the backup software configuration to begin storing recovery copies. A DD400 is deployed as the online, on-site recovery repository in the following configuration:

A DD400 is added as on-site backup storage system, taking the place of any disk staging devices and the on site tape library. Backups are directed to the DD400 for improved backup performance. As a DD400 holds many months of backups, recoveries are also performed from the DD400, speeding up the recovery process. Backups may then be copied to the tape library (or network vaulted to an offsite DD400 and then cloned to tape), providing tape copies for off-site storage for disaster recovery or long-term archive purposes.

This configuration lowers the overall cost of the backup and recovery infrastructure, with lower library, handling and operational costs. The tape library can still be dedicated to provide disaster recovery protection. As the DD400 is the primary backup target and recovery source, the demands on the tape library is reduced. A smaller tape library can fulfill the role of a "clone tape library," which streamlines the backup and recovery process and infrastructure.

Further use of restorers in a remotely mirrored configuration via DD Replicator connecting them, as shown above, offers additional benefits in automation, reliability, disaster preparedness and time to restore by enabling fully tapeless data protection.

Easier to Administer than Standard ATA RAID

When compared with ATA RAID-based alternatives, consider these differences. Notably, a DD400 offers dramatic resilience through verifiable recoverability and self-healing data. In addition, a DD400 has significant ease-of-administration and cost advantages:

- **Operational simplicity.** To store the amount of data in a DD400's 15 disks, standard RAID 5 implementations would require about 300 disks, and mirroring would require more than 500 disks (depending on implementation, file system overhead, and so on). More disks mean more rack space, more power, more administration, and more hassle. If disks fail at

a consistent rate, that's an order of magnitude more failures manage.

- **Ease of total system management.** Most systems today are external block arrays, so they are that much more difficult to administer end-to-end compared to the integrated approach of a DD400.
- **Cost.** Disks are the largest cost element of ATA RAID. Fewer disks mean lower cost.



Summary

The DD400 Enterprise Series deliver all of the advantages of disk-based backup storage with the economy of a tape based storage architecture: faster and easier backups and restores, simultaneous access to allow multiple, concurrent backups and restores or multiple backup servers, and faster and easier ways to create off-site replicas of backup data, either using tape or network vaulting.

However, the DD OS and DD400 benefits go far beyond general purpose RAID disk arrays. The DD400 is specially designed as a backup storage appliance, delivering:

- **Economy of tape** through Global Compression for a lowered TCO
- **High performance backups and restores** to meet shrinking backup window and recovery time requirements
- **Data integrity; verifiable recoverability and data self-healing** to ensure reliable recoveries
- **Simple and seamless integration with backup software** to support backup software standardization and leverage existing investments
- **High efficiency network vaulting. Fast and low-cost replication to remote data centers** to free the primary data center from tape handling

Technical Specifications

SOFTWARE

Software Features

Data Domain OS (DD OS) 3.0

Global and local compression, integrated dual parity RAID (DDRAID), Data Invulnerability Architecture™, telnet, ftp, ssh, email alerts, command line management interface, GUI, end-to-end verification (ongoing), scheduled capacity reclamation (clean), SNMP, Data Domain Replicator (add-on option)

File Protocol Support

NFS V3 over TCP, CFIS, NDMP v2

HARDWARE PLATFORM

Normal Operating Current

3U 19-inch rack-mount enclosure, hot-plug disks, redundant fans, N+1 power supplies, (2) 10/100/1000 Ethernet port (optional dual add-on GB Ethernet, copper or fiber), serial port

Maximum Current at Power Up

(for 115VAC/230VAC) DD410/430: 3.0/1.5A; DD460: 3.5/1.8A

System Thermal Rating

(for 115VAC/230VAC) DD410/430: 4.5/2.3A; DD460: 5.0/2.5A

System Weight

DD410/430: 1370 BTU / hr.; DD460: 1570 BTU / hr.

System Dimensions (WxDxH)

DD410/430: 60 lbs (28kg); DD460: 80 lbs (39.3kg)

Operating Temperature & Humidity

19" x 26" x 5.25" (4.83 cm x 66 cm x 13.3 cm)

Non-operating Temperature

5C-35C @ 5%-90% RH, noncondensing

Minimum Clearances

(40) C - 65C

Front, w/Bezel Closed

1" (2.5 cm)

Front, w/Bezel Open

5" (12.7 cm)

Rear

5" (12.7 cm)

Operating Acoustic Noise

Max 75 dbA, at rear of unit when all drives seek simultaneously

REGULATORY APPROVALS

Safety

UL1950

Emissions

FCC Class A

Immunity

EN 55024: 1

Product Family

	DD410	DD430	DD460
Capacity: Raw	8x160 GB	8x400 GB	15x400 GB
Capacity: Standard(1)	15 TB	42 TB	83 TB
Capacity: Redundant(2)	55 TB	101 TB	233 TB
Throughput maximum	160 GB/hr	220 GB/hr	290 GB/hr
Consider for primary data of this size(3)	Up To 1 TB	Up To 2.5 TB	Up To 5 TB

1. Mix of typical enterprise data (file systems, databases, mail, developer files), full backup weekly, incremental backup daily, to system capacity
2. Mix of typical enterprise data (file systems, databases, mail, developer files), full backup daily, to system capacity
3. Estimates based on conventional use. Please contact Data Domain Sales for site analysis.

References

Content-based segmentation

- Udi Manber. [Finding similar files in a large file system](#). In *Proceedings of the USENIX Winter 1994 Technical Conference*, pages 1–10, San Francisco, CA, USA, 17–21. 1994.

Use of a fingerprint as a hash to compare two segments

- M. Rabin. [Fingerprinting by random polynomials](#). Report TR1581, Center for Research in Computing Technology, Harvard University, 1981.
- A. Broder. [Some applications of Rabin's fingerprinting method](#). In R. Capocelli, A. De Santis and U. Vaccaro, editors, *Sequences II: Methods in Communications, Security, and Computer Science*, Springer Verlag, 1993.

High-speed indexing and verification: Data Domain restorers and replicators (40-80MB/sec NFS throughput)

Related research prototypes (an order of magnitude slower)

- A. Muthitacharoen, B. Chen, and D. Mazieres. "A low-bandwidth network file system." ACM SOSP Conference, 2001.
- S. Quinlan and S. Dorward, "Venti: a new approach to archival storage", USENIX FAST conference, 2002.

About Data Domain

Data Domain is the leading provider of Capacity Optimized Storage systems. Through advanced Capacity Optimization, data invulnerability and replication technologies, Data Domain systems enable storage of backup/recovery data through traditional backup software to be simpler, smaller, faster, and more reliable.

Visit us online at www.datadomain.com.



Data Domain DD400 Enterprise Series White Paper - 2005



3400 Hillview Ave.
Bldg 3, 2nd Floor
Palo Alto, CA 94304
877.622.2587
sales@datadomain.com