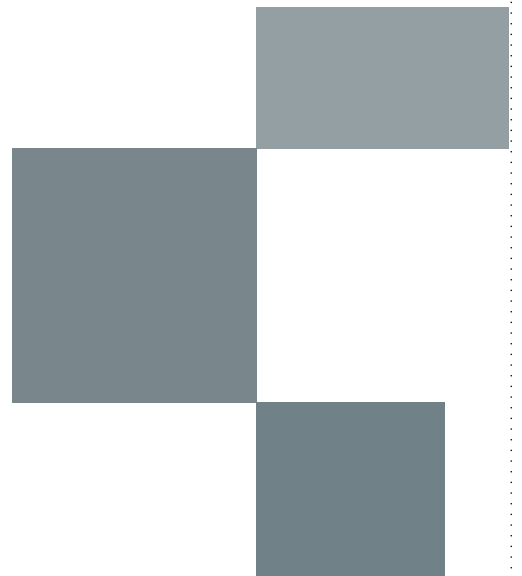




Capacity  
Optimized  
Storage

## Eliminating Tape and Trucks from the Disaster Recovery Process

Cost-Effective WAN Vaulting Solutions Deliver Automated, Reliable Recovery of Critical Data





THE TROUBLE WITH TAPE: MANUAL AND UNRELIABLE ..... 3

DARWINIAN DISASTER RECOVERY: EVOLVING BEYOND TAPE..... 4

CAPACITY OPTIMIZED STORAGE: REVOLUTIONIZING STORAGE ECONOMICS ..... 5

THE DATA DOMAIN DD400 ENTERPRISE SERIES: WAN VAULTING FOR ENTERPRISE BACKUP ..... 7

SUMMARY ..... 9

ABOUT DATA DOMAIN ..... 9

## The Trouble with Tape: Manual and Unreliable

Since the dawn of enterprise computing, disaster recovery has been of key concern – how can companies that rely on computers continue to operate after any sort of business disruption? Despite constant overall technologic change, data storage options for disaster recovery have changed remarkably little in the last three decades. On one end of the spectrum, for extremely high-value data, “hot standby” duplicate environments are the top choice of enterprises for which seamless business continuity is absolutely essential; these environments house exact replicas of the hardware, software and data contained in main data centers, and are ready to be called to duty at seconds’ notice. Data storage in these failover environments requires extraordinary resources to be kept synchronized in real-time: multiple, ‘round-the-clock, high-speed connections (typically OC3 or dark fiber) with the “base” data centers.

Relegated to the low end of the disaster recovery spectrum, unfortunately, is virtually “everyone else” – every enterprise, remote office, department and small or medium business that cannot afford dedicated high-speed connections to remote disk-based storage. In reality, the default data transport choice for these organizations is magnetic tape, a “one size fits all” medium that, unfortunately, is unreliable, extremely cumbersome and highly susceptible to human error.

Gartner’s estimation of tape reliability – that one in 10 recovery images on tape is unrecoverable – is optimistic.

## Tape Backup: Low-cost, low reliability, high effort

Although it has been in existence for half a century, tape has, until recently, remained the most cost-effective method for storing and retrieving data. Tape offers adequate data protection at a low cost – its main benefit. But unfortunately many companies find that the fastest way to transport data stored on tape is not over bandwidth, but by truck – a reality that, in the aftermath of events such as the terrorist attacks of 2001 and more recently, Hurricane Katrina, is simply unacceptable and at times unattainable.

Events such as these make clear that Gartner’s estimation of tape reliability – that one in 10 recovery images on tape is unrecoverable – is optimistic. Even if it were correct, a 10 percent failure rate is by far the worst operational dependency in enterprise IT. Worse still, the failure rate can only be measured when there’s a problem – an unfortunate time to discover it. To compensate, some IT organizations make multiple copies of every tape backup, despite the skyrocketing effect this practice has on the “low cost” of tape-based disaster recovery.

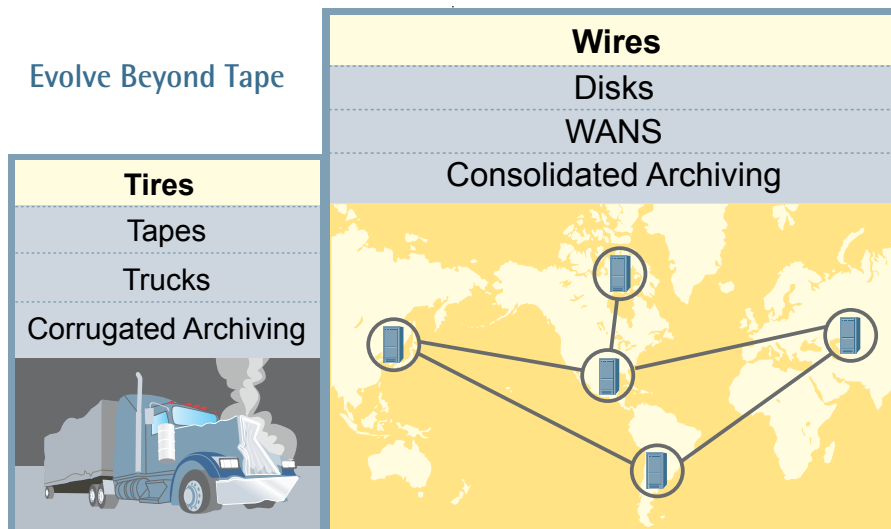
Additionally, the processes required to copy and then restore production data from backup tapes rely heavily on human intervention – the corollary of which is inherent delay and potential for error. In a global organization with multiple backup storage sites, this labor-intensive process also creates a risk for data theft, a costly public relations catastrophe that reduces customer confidence. In sum, tape backup

unequivocally translates into large amounts of effort and risk.

### Capacity Optimized Storage: A cost-effective, automated, reliable alternative

Capacity Optimized Storage (COS) is an emerging category of disk-based backup storage solutions that directly addresses the data volume challenge. For the first time in decades, it presents a viable alternative to tape backup, offering superior price-performance and far more responsive disaster recovery capabilities. Capacity Optimized Storage is enabled by deduplication technology, which massively reduces data (by more than 20x) down to its smallest possible size, into an amount of bytes that can be stored or easily transferred over the network to a disk system in the disaster recovery site, and readily be retrieved should the need arise.

This whitepaper explores the limitations of tape and traditional disk-based approaches to disaster recovery. It highlights the advantages of deduplication technology in disk-to-disk backup storage and wide area network (WAN) vaulting. It also discusses related applications for COS such as remote office backup and tape consolidation among distributed data centers.



## Darwinian Disaster Recovery: Evolving Beyond Tape

The popularity of disk backup has literally exploded since the turn of the millennium. This fact springs from a basic operational irony: tape is simply too fast. In order to keep tape drives spinning, data must be fed at very high speeds. Since most backup clients can't feed the tape drives fast enough, IT organizations have two choices:

- Multiple client streams can interleave or multiplex across the same tape, which makes recovery slow and error-prone;
- Backups can be cached first to an isolated disk store to ensure they can stream to tape later

Caching has the added benefit that recoveries from the cache disk are random access and fast. But backup data volumes are too large to be kept in the cache.

Alternatively, tape libraries generally house two to three months of full and incremental backups – which translates into storing five to ten times as much data as is being backed up in the first place. Therefore, for cost reasons, as much backup as ever is still being performed

with tape. Backup to disk has not had a major impact on traditional tape-based backup practices. It does improve short-term data retention – but disaster recovery is still a truck-based affair.

### Traditional disk-based disaster recovery: Cost-prohibitive & process heavy

In addition to caching high-priority data to backup disk, full data replication to disk is a second potential strategy for disaster recovery. While easy to carry out day to day, it has two major pitfalls when considered for disaster recovery: high cost and process augmentation. Copying production data across a WAN to disk is too expensive for most applications.

Even block delta replication, considered a best-of-class alternative for minimizing WAN bandwidth consumption, often entails transporting data volume of five to 10 percent of the protected quantity every day. Given the high costs of WAN communications, this approach is often too expensive.

A WAN/disk-based approach also requires infrastructure and process augmentation, since replication augments backup software with its own approach to data capture and recovery. Even when replicated, data is still backed up. Keeping months of

versions of data is too expensive on primary storage, compared to tape or disk backup. In addition, replication is typically array-based, so it does not consolidate all server data.

### Capacity Optimized Storage: Revolutionizing Storage Economics

Deduplication is a new technology at the heart of

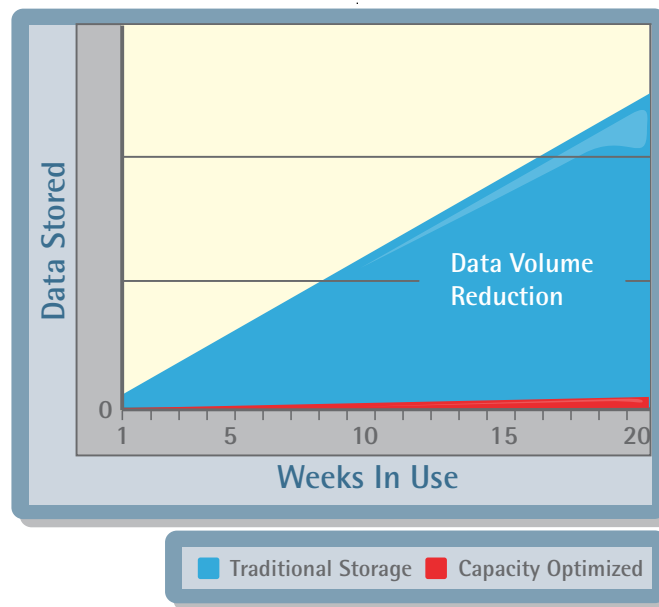
Capacity Optimized Storage. Deduplication is designed to massively reduce data down to its raw essentials, thus optimizing capacity and making full disk backup and WAN Vaulting economically and operationally feasible. Unlike traditional compression, which typically reduces data by half, deduplication can condense standard business backup data down to a twentieth or

less of its original size. This is achieved by breaking the data into a small number of fundamental sequences which, when replicated, can be used to rebuild the original data.

The analyst firm Taneja Group explains:

“Capacity optimization (CO) is being used both in storage devices and also networking devices to build

Reduce Data Volume for Network-efficient WAN Vaulting



much more cost-effective systems. For example, a storage device that utilizes CO can store 20x as much data as a standard one at the same cost. A capacity optimized network can transmit 20x as much data as a non-optimized one again all for the same price. This order of magnitude cost reduction in capacity optimized technologies vs. standard ones is ensuring that all future system designs will at some point implement this technology.”

“The primary emerging market utilizing CO today is the data protection marketplace. In this market, corporate data has traditionally been stored on tape rather than disk. Although disk is widely recognized to keep data in a manner that is safer and more accessible than tape, until now it had been too expensive to use for longer-term data protection, due to cost differentials making tape about 1/20th the cost of disk storage solutions. By utilizing CO, a new generation of solutions are able to provide a massive savings over regular disk based storage, bringing them into price equivalence with more traditional

tape solutions.”<sup>1</sup>

### End-to-end Tapeless Data Protection: Off-site vaulting for DR

Capacity Optimized Storage is ideal for disaster recovery applications. By reducing data volume, deduplication enables efficient bandwidth utilization for automatic WAN vaulting of backup data to disaster recovery sites. As the term implies, WAN vaulting delivers cost-effective offsite storage capacity similar to the physical vault where backup tapes are stored, but with immediate accessibility via the network. It is the best industry approach for 90 percent of enterprise applications and data.

According to research firm, Enterprise Strategy Group, 30 percent of corporate information is in major data centers, but another 30 percent resides in remote offices. (The remaining 40 percent is stored on desktop and laptop computers.) Today, most remote office data is backed up with tape, often

### Off-site Vaulting for DR and Tape Consolidation



<sup>1</sup>“Introduction to Capacity Optimization,” The Taneja Group, May 2005

through the efforts of a local employee who may not be an experienced IT system administrator. As a result, tape backup is often executed sporadically and/or improperly, increasing its propensity to fail and decreasing predictability.

By vaulting backup data across wide area networks, this problem is solved easily and effectively, at the absolute minimum WAN cost. The local tape autoloader can simply be replaced with capacity optimized disk-based storage to provide local recoverability that is highly reliable. Backups can be replicated to a central hub site or sites, where a larger system can store the replica backup data for several remote offices.

Once the data has been written at the hub, it can then be moved to physical tape as required for archiving. Administration of all required backup operations can all be done remotely over the network.

### Streamlined, cost-effective tape consolidation

Because tape is a highly manual medium, the more its use can be centralized and consolidated to one location and disk used everywhere else, the more likely that operational recovery and archiving will be done correctly.

With WAN vaulting, many topologies are easy to create that enable successful consolidated tape archiving. These include:

- In a simple case of two offices, only selected recovery images may be sent to a replica system.

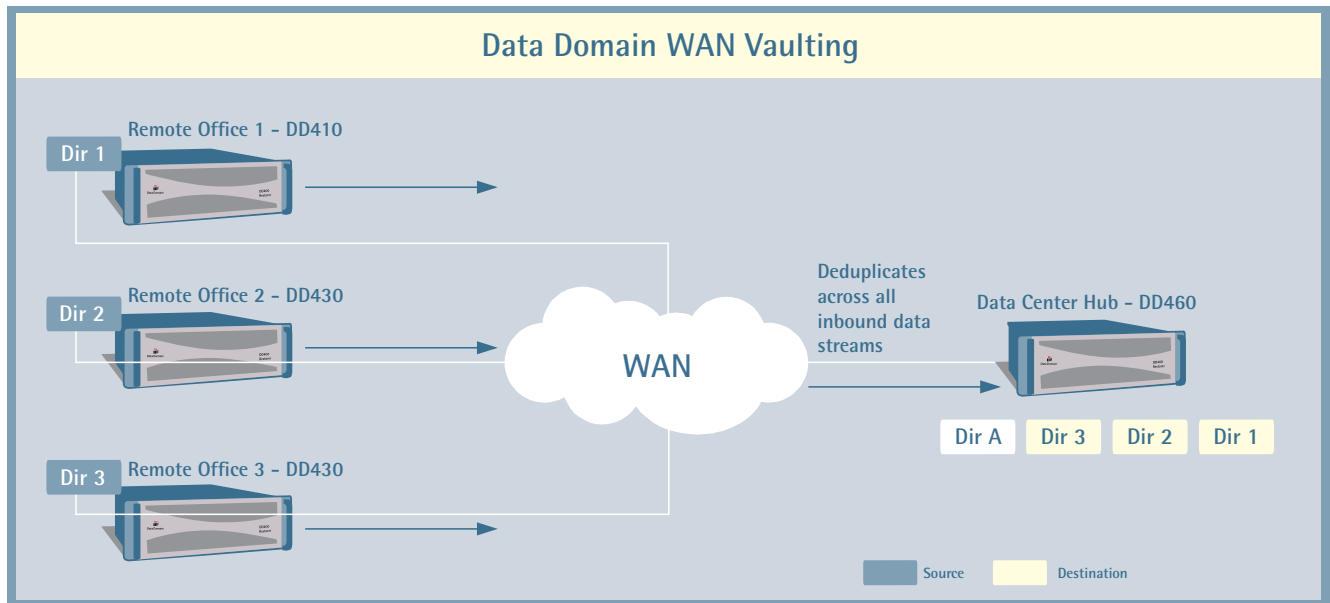
For example, if only Microsoft Exchange backups need to be archived, but file share backups do not, the system can be set it up such that only the Exchange backups are replicated to a safe site, where they are then duplicated to tape.

- In a more complex two-site case, there may be backups in both sites, but tape is centralized at one. Both nodes replicate to each other. This allows either site to have all backup data, and duplication to tape to be managed in either site.
- As discussed above, in a many-to-one replication scenario, there may be many small sites and one hub site that includes tape. Tape consolidation is a natural outgrowth of this architecture.

## The Data Domain DD400 Enterprise Series: WAN vaulting for enterprise backup

The Data Domain DD400 series of restorers is the industry's leading family of Capacity Optimized Storage restorers for backup/recovery storage with WAN vaulting for enterprise backup. Data Domain's replication software option, DD Replicator, enables network-efficient WAN vaulting for disaster recovery, remote office data protection and tape consolidation. All DD400s are designed with Data Domain's deduplication technology, Global Compression™. Key DD400 benefits include:

- Less than \$1/GB for disk backup storage
- 99% bandwidth reduction for WAN Vaulting
- 20x data reduction for capacity optimization



## Data Domain's deduplication enables world-class Capacity Optimized Storage solutions

The DD400 Series offers an average 20x data reduction for enterprise recovery images, enabling cost-efficient retention on disk for high-speed and more reliable recoveries. As a result, it enables a 99 percent bandwidth reduction in WAN-based data replication, making WAN vaulting for enterprise backup possible.

The DD400 Series offers capacities from 15 TB to more than 200 TB of usable storage per system for a typical enterprise data set and backup policy. With its high performance system architecture, the DD400 Series offers up to 290 GB/hour of throughput, storing data at single or multiple sites.

The high capacity and throughput of the DD400 Enterprise Series lets users back up massive amounts of data to disk under tight time constraints; they can

compress and store on average four to five months of backup data onsite and online, simplifying the backup and restore process and virtually eliminating reliance on tape storage. In addition, storage administrators no longer have to waste time monitoring slow tape backups and re-running backups each time they fail. Users can replicate their data to remote locations, further reducing reliance on backup tape and offsite storage vendors.

These reductions in total cost of ownership (TCO) are coupled with dramatic hard-cost savings; the DD400 Enterprise Series delivers unprecedented compression capabilities, bringing the cost of disk-based backup to less than \$1/GB.

## Easily inserts into existing enterprise backup infrastructure

Data Domain restorers are qualified with all leading enterprise backup software – no infrastructure change is required for either data center or distributed office data protection. As a result, the

DD400 Enterprise Series can be used to protect the complete range of enterprise or developer application data including:

- Oracle and other databases
- Exchange and .PST files
- Home directories
- Developer files, e.g. ClearCase, CVS
- Enterprise applications such as CRM and ERP

## Summary

Disaster recovery and remote office replication are top-of-mind issues for IT professionals, with recent disasters such as Hurricane Katrina underscoring the fallibility of the tape-based backup strategies commonly used by all but the largest enterprises. Tape-based backup is woefully inadequate for disaster recovery situations because it involves too much human intervention, and the management of too much data – two conditions that almost inevitably lead to error and delay.

Data Domain Capacity Optimized Storage solutions dramatically reduce storage and network costs, leveraging current infrastructure. They offer:

- **Lower cost of operations** by reducing the amount of data sent over a WAN by 99 percent and more, compared to replicating any other disk-based backup system for standard backup software. This is enabled by Data Domain's patented deduplication technology and advanced

compression algorithms, which allow compression rates of 20x to be readily achieved.

- **Improved recovery time** vs. tape. Files can be recovered locally from disk or over the network from a replica system, free of tape-finding or mounting delays. Given the huge human element in tape-based systems – and associated fallibility – Data Domain removes one of the largest obstacles in attaining seamless business continuity.

By replacing tape-based backup systems, Data Domain dramatically improves the speed and efficiency with

which disaster recovery can occur – and eliminate tape, trucks and unnecessary time, once and for all.

Data Domain Capacity Optimized Storage solutions dramatically reduce storage and network costs, leveraging current infrastructure.

## About Data Domain

Data Domain is the leading provider of Capacity Optimized Storage (COS) solutions, enabling reliable, cost-effective data protection for the enterprise. More than 300 companies worldwide use Data Domain's award-winning disk-based backup and recovery storage systems. Data Domain's Global Compression, data invulnerability and replication technologies offer breakthrough compression rates reducing the cost of disk-based backup and simplifying data recovery. Founded in 2001, Data Domain is a privately held company. For more information, visit Data Domain's web site at [www.datadomain.com](http://www.datadomain.com). Data Domain is headquartered at 3400 Hillview Avenue, Palo Alto, California 94304 and can be contacted by phone at 1-877-622-2587 or e-mail at [sales@datadomain.com](mailto:sales@datadomain.com).

# Eliminating Tape, Trucks and Time from the Disaster Recovery Process

3400 Hillview Ave.  
Palo Alto, CA 94304  
877.622.2587  
sales@datadomain.com