The ROI and TCO Benefits of Data Deduplication for Data Protection in the Enterprise

A Focus White Paper

June 08

FOCUS consulting

Sponsored by: datadomain

Table of Contents

The ROI and TCO Benefits of Data Deduplication in the Enterprise

Executive Summary

Unreliable backups and recovery times that are too long have long plagued IT organizations. In many ways, data protection has unsettling similarities to insurance: it's required, it's expensive, and one hopes never to need it. And, like insurance, one's objective is to have enough but not too much. To that extent, there is an inevitable cost/benefit business case analysis to any data protection purchase.

A tape-only data protection scheme is no longer a best practice. Tape simply does not deliver the necessary backup reliability, rapid recovery, or recovery point precision needed by most organizations. Virtual tape libraries (VTLs) reduce the mechanical aspects of tape backup, but are really an intermediate step in what is otherwise an unchanged process. To address the deficiency of tape and VTLs, many organizations are examining deduplication storage systems. These systems eliminate the mechanical reliability problems of tape, reduce the need for media handling and operator intervention, and increase the overall efficiency and reliability of data protection operations.

Even though tape is being phased out for day-to-day operations, it is still used by many organizations for long-term data archive. Thus, there is no denying that deduplication storage systems are another layer of infrastructure. The question that must be answered is whether the benefit of this additional layer is a net savings gain and operational improvement over tape alone.

In evaluating the business case for data protection products, four primary criteria are used:

- Purchases avoided
- Direct savings
- Labor savings
- Total cost of ownership (TCO)

This paper examines and quantifies the costs and benefits of backup with deduplication storage as strategic assets for data protection. Three organizations in different industries are discussed as case studies. Each company was examined based on these four criteria. For each of these companies, operational issues were the driving force behind the purchases — the IT managers involved in the decisions wanted more reliable backups with less operator intervention. However, each company's results show a substantial cost savings as well.

Deduplication, or the process of eliminating redundant data that is written to a storage device, is an enabling technology that has gained significant traction over the past several years. Deduplication storage for data protection delivers three primary benefits:

- Shortens backup times through disk-based storage and deduplication rather than tape, allowing backups to fit within required backup windows
- Makes data replication over a WAN realistic by reducing the amount of data transferred
- Permits data to be retained within the data center for longer periods due to smaller data sizes.

In the case studies presented here, the three-year ROI ranged from 16% to 62%, while the payback ranged from just three months up to 25 months. In all cases, the companies realized the business and operational benefits they were seeking, and the cost savings provided both the financial justification and the added bonus of proving the system's business value over time.

Introduction

Meeting the Challenges of Backup and Recovery

Exponential data growth, changing regulatory requirements, and increasingly complex IT infrastructure all complicate data managers' data protection schemes. Although the past days of nightly backup to tape are arguably less complex than today's 24 x 7 availability backup environment, no one would refer to them as "the good old days." Given the mechanical nature of tape, operational discontinuities are inevitable. While most organizations do not realistically expect to eliminate tape backup entirely, most are interested in minimizing its use and corresponding operational problems.

Data protection best practices are evolving at a steady pace. Recovery time objectives (RTO) continue to decrease while the precision of the recovery point objective (RPO) increases. In other words, IT managers must be able to recover from a given failure quicker and with less data loss. The time needed to find, mount, and search tape media is not improving fast enough to keep pace with the changing RTO and RPO requirements of most organizations.

The time needed to find, mount, and search tape media is not improving fast enough to keep pace with the changing RTO and RPO requirements of most organizations.

It is not uncommon for organizations to routinely exceed their backup window or even have a backup window that consumes most of the day. Such long backup operations leave little margin for error and any disruption can place at least some of the data at risk of loss. Such operations also mean that a guaranteed RPO of anything less than 24 hours cannot be met. To make things worse, tape backup schemes (i.e., grandfather-father-son) require that the same data be backed up over and over again. Of course, incremental backups reduce backup times and the number of data images, but at the cost of unacceptably long recoveries.

The backup window is further challenged by constant and inevitable data growth. Because tape backup is linearly proportional to the data volume, increases in volume will extend the backup window. For example, a 20% increase in volume will cause a 20% increase in backup time. The only way for organizations to respond when using tape is to either upgrade to a faster technology or deploy more tape devices. Either way, it is a significant investment that does not fundamentally change the data protection service level delivered to the organization.

More and more organizations are grappling with the difficulty of managing remote office operations. Because of mergers and global business expansion, IT organizations are confronting far-flung operations. In many cases, these remote offices do not have professional IT staff, but still have the same data protection mandates. Central control of remote backup is essential to maintaining the data integrity demanded by the business.

The most challenging factor for the data center manager remains the human element. The more complicated the backup operations become, the more skilled people are needed. These people can be difficult to recruit, and human costs continue to rise. So, even if qualified individuals can be found, management may be reluctant to hire them. The mantra is "do more with less (or the same)."

Evolving Technology

Tape technology is by no means static. However, it has been many years since the industry has seen a fundamental breakthrough in this area. To be sure, tape drives continue to get faster and tapes increase in capacity. Even so, this is simply more-of-the-same-but-faster.

Automating tape technology also has its limits. Automated robots make picking and mounting tapes faster, but offsite archive and retrieval cannot be automated. Tape media must still be handled manually inside and outside the data center. The extent to which labor can be reduced represents the greatest potential savings to organizations. Transportation of tapes outside the data center also introduces a major security risk, through the possibility of lost or stolen tapes.

Although long-term storage of data on disk is becoming more economical, tape remains the predominant technology for data archive. Storing data on tape in an offsite archive is economical, but makes recovery certainty problematic after long periods. Data formats, tape formats, and tape devices change gradually over time. This creeping obsolescence may mean that organizations attempting to recover data five or more years after its creation will unexpectedly face a lengthy and costly effort.

In the 2004/2005 time frame, disk-based backup systems were a curiosity within the industry. They have since achieved mainstream status. Our 2007 research indicates that more than 50% of Global 2000 organizations utilized some form of backup to disk. We expect this adoption to grow to 80% by 2013. Organizations that have adopted backup appliances did so most commonly for the following reasons:

- To gain higher reliability and certainty of backup job success
- To reduce labor associated with tape handling
- To obtain longer retention of backed up data within the data center
- To consolidate data from remote offices for centralized backup

Changing the Game: Deduplication

Backup to disk alone is not a game-changing development. Backup to disk has been utilized for many years, with products such as Tivoli Storage Manager, to stage the data for tape offload. Data deduplication, unique to backup appliances and pioneered by Data Domain, can fundamentally change the way organizations protect backup and nearline data. Deduplication changes the repetitive backup practice of tape, with only unique, new data written to disk. However, the deduplicated backup image does not carry the restore penalty associated with incremental backups because the entire image is still available on the same device. Furthermore, with an average of 20:1 deduplication ratio commonly achieved with backup appliances, backup data can be retained economically in the data center for long periods of time. This reduces the odds that a data element must be retrieved from the vault. Both of these factors can significantly improve the RTO.

Justification

Business case drivers such as more reliable backups and faster data restores indirectly address cost reduction. The impetus to buy a backup appliance will almost always come from the IT group and not an

end-user business unit. Nevertheless, purchase approval from non-IT decision makers is almost always required.

Non-IT decision makers can intuitively understand the value of longer data retention, faster recovery, and reduced labor. However, these are the common justifiers for many technologies. Consequently, such "soft promises" may not be sufficient to justify a capital outlay; technology benefit promises of the past have not always been kept.

Although cost savings are generally not the initial reason to consider moving to disk backup and deduplication, financial justification is almost always a prerequisite. With the potential cost savings of disk and deduplication, the justification statement becomes, "We can get all of these business benefits and save money." That's a compelling argument.

Although cost savings are generally not the initial reason to consider moving to disk backup and deduplication, financial justification is almost always a prerequisite. With the potential cost savings of disk and deduplication, the justification statement becomes, "We can get all of these business benefits and save money." That's a compelling argument.

Unfortunately, promises of cost savings have an even more checkered history than promises of benefits. Savvy managers want to see some historical proof to back up the claims. This white paper discusses case histories from real-world scenarios of three Data Domain enterprise customers, documenting the actual savings experienced by these customers, in different industries, with their equipment installed for as long as three years.

In past studies, avoiding the need to upgrade tape infrastructure was the key factor in the financial justification, one that offset the cost of a new Data Domain system. In these three case studies, however, avoiding a new tape purchase was a minor consideration. Instead, cost savings for these enterprise customers were realized primarily through reduced administrative labor.

ROI/TCO Analysis Methodology

The ROI and TCO information presented in this white paper is based on a financial analysis conducted by Focus analysts of three enterprise-class Data Domain customers with at least a one-year history with Data Domain products. The actual product usage history by customer ranged from one to three years. Historical financial data was used whenever possible, with future numbers based on financial projections from the customers, based on past experience.

The data from the customers was entered into a customized ROI/TCO calculator created by Focus. This unique tool takes all of the data into consideration, and calculates both ROI and TCO. All of the numerical information and charts in this report were created using the Focus tool.

The Business Case

In presenting these case studies, this paper discusses several financial-related terms: return on investment (ROI), total cost of ownership (TCO), payback, labor savings, and cost avoidance. A definition of each term follows:

TCO (total cost of ownership)

A TCO model establishes a fully loaded, total cost of a project over time. Decisions are made by comparing the TCO of one approach to the TCO of another. TCO is a cumulative number, over some period of years (typically three for IT), and incorporates the changes in costs and benefits over that period (e.g., due to data and storage growth). TCO includes capital acquisitions, maintenance, and operational costs, and should include both cost components that are *direc*t (e.g., hardware and software acquisition, salary costs of full-time employees) and *indirect* (which are often difficult to quantify, such as the cost of waiting for a file to be restored). The TCO categories used in this paper are ***Hardware, Software, Support, Supplies and Services.*** Salaries generally are based on a 30% burden rate, to cover insurance, benefits, etc.

ROI (return on investment)

ROI is a measure of the financial return on an investment over a specified period of years (typically three for IT), represented as a percentage. A minimum ROI may be required by corporate finance departments in order to get approval on a project/acquisition.

Payback

Payback period is the amount of time it takes for a project to pay for itself or break even, such that the cash flow becomes cash positive for all aspects of the project.

Total savings

Total savings is the amount of both direct and indirect dollar benefits resulting from the project.

Net savings

Net savings is the net amount saved over a given time, calculated by subtracting the costs for that time period, from the total savings for that time period.

Direct savings

When the project results in a direct cost reduction, where cash outflow is reduced, these reductions are direct savings. Significant direct savings described by users in this paper's case studies include:

- Supplies and services — These types of direct savings involve a reduction in the total cost of tape media and the services to transport and maintain those tapes offsite. For users with a large number of tapes, these savings alone can be staggering.
- Cost avoidance in hardware — These savings are the result of eliminating the need to purchase additional tape hardware to complete backups within the available backup window, as well as for performing tape backup in remote sites. For users already up against the window, or for users eliminating tape in remote sites, this can be the largest percentage of savings.

Indirect savings

When implementing a project can save time (for IT staff or end users), the result is considered indirect savings. Indirect savings include:

- Cost avoidance in labor — Labor cost avoidance is time saved by backup administrators, systems administrators, or end users as a result of implementing the project. This savings would allow the user the choice of either spending time on other projects or potentially reducing headcount (of full-time equivalents or FTEs). For purposes of this paper, this category is calculated as a cost reduction.

In the three studies presented here, the three-year ROI ranged from 16% to 62%, while the payback ranged from just three months up to 25 months. In all cases, the customers realized the business and operational benefits they were seeking, and the cost savings provided both the financial justification and the added bonus of proving the system's business value over time.

Case Study #1: European Geospatial Data Customer

Managing geospatial files presents a number of unique data protection challenges. These files tend to be very large, yet remain static throughout their useful life. They are also infrequently accessed, but when access is required, it must be prompt and reliable.

To date, tape has been the primary method of protecting geospatial data, largely because it was the only game in town. However, static files do not benefit from the same daily-weekly-monthly backup routine as might other data. Once the image is captured, it should be good for its useful life. Because of this factor, it may seem tempting to back up the data a few additional times for disaster recovery and offsite vaulting. However, having the image of an entire disk array spread across potentially dozens or hundreds of tapes would make recovery time totally unacceptable. Even routine file restore operations would be measured in hours or days.

Consequently, geospatial organizations are forced to back up the data using common grandfather-father-son methodologies. Managing all of the data images results in gross inefficiency of human resources due to labor spent on tasks not strictly needed to protect the data. Moreover, a 10% to 20% nightly average tape job failure rate is well documented within the industry.

Such was the case with the first Data Domain customer case study. This European-based organization produces land survey data on a nationwide basis, with some maps dating back more than 200 years. Current mapping technology uses high-resolution digital technology creating files of 700 MB or more each. The organization experienced a 15% failure rate for nightly backup jobs, mostly due to media failures. Moreover, the organization frequently saw the backup window extend into daily operations.

To solve this tape processing problem, the organization examined Data Domain's appliance series. The organization became convinced that Data Domain's operating architecture would allow it gradually to eliminate its reliance on tape for daily backups. This operational change would reduce the labor needed to manage tapes and forestall any upgrade to the tape environment. Data Domain's deduplication technology would reduce backup times by eliminating the need to back up a static data file over and over.

In conjunction with these benefits, the organization needed to examine the cost justification for purchasing the necessary Data Domain solution. The organization began by installing two Data Domain

DD560 deduplication storage systems in April 2007, placing one in the data center and one in a remote site for offsite replication.

Prior to installation of the DD560 systems, the organization had 80 TB of data and managed 4,300 tapes. The customer experienced direct savings in services and supplies, as well as indirect savings on labor and hardware purchases that were avoided. Figure 1 provides a detailed breakdown of these cost savings.

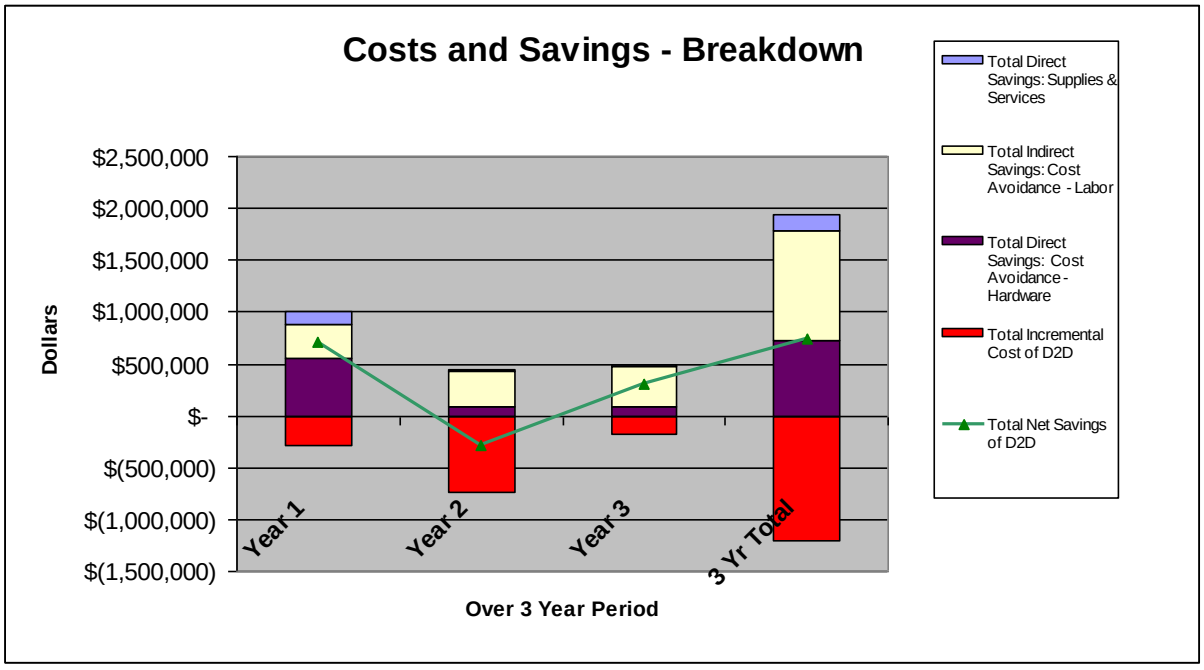

Figure 1: Breakdown of costs and savings for geospatial data customer

As Figure 1 indicates, the organization saw an immediate cost savings in the nine months following installation. (Net Savings are indicated by the green line.) After the successful implementation of the DD560 systems, the organization implemented a DD580, two DD565s, and three ES20 expansion shelves. These systems were put into service in January 2008. Total net savings over three years exceed $730,000. Net savings are calculated by subtracting the Incremental Cost of disk deduplication storage (shown in red) from the total savings (the sum of Total Indirect Savings (labor cost avoidance) and Total Direct Savings (supplies and services and hardware cost avoidance).

Figure 2 illustrates the breakdown of the various savings. As the customer expected, substantial labor was saved by managing far fewer tapes and by not having to reinitiate failed backup jobs, among other things. Labor savings accounted for more than half of the total savings (54%).Given that “Efficient backup was the primary reason for purchase” (with additional security by means of WAN data vaulting as an important feature), the customer was “very happy with the results.”

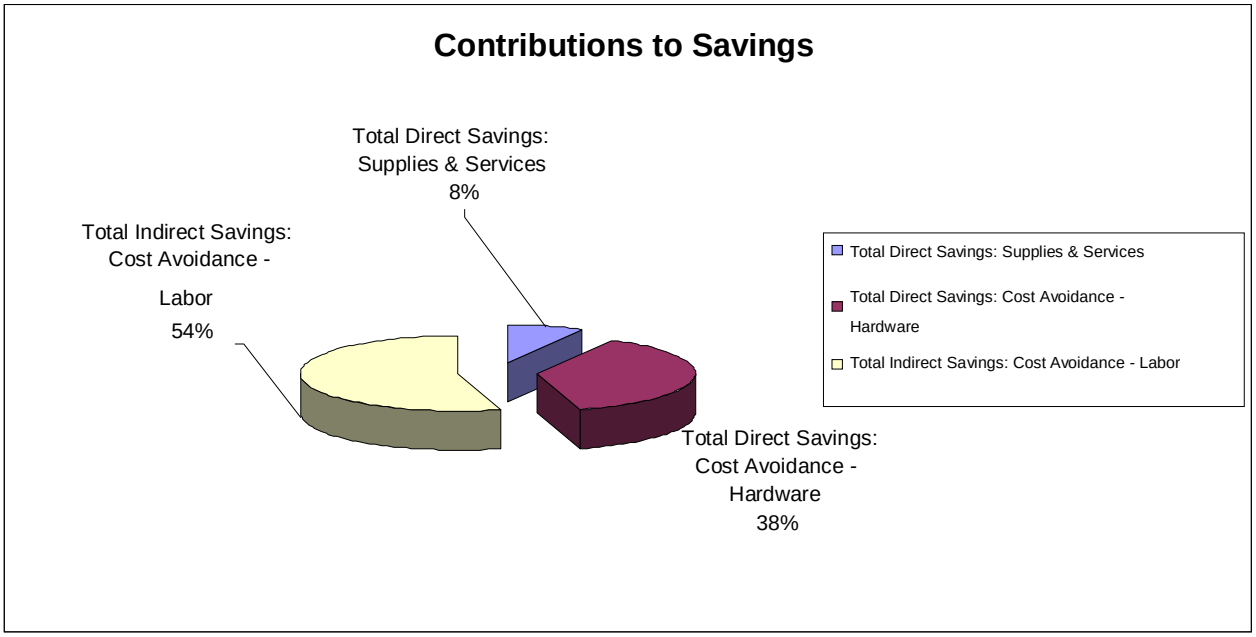

Figure 2: Contribution to savings by category for geospatial data customer

Cost savings is a key justifier of any new technology. However, the new system must also prove superior to other viable solutions. In this organization's case, the alternative was to upgrade the existing tape infrastructure. The customer had three tape libraries with LTO-1 technology installed. The IT staff considered adding a fourth library and upgrading all systems to LTO-4. The cost for this upgrade was quoted as $368,280. Over three years, the customer was able to avoid purchasing $731,808 in tape system upgrades. Although the customer understood the risk of using an older tape technology, the issue was somewhat moot, as the intention is to eliminate the use of tape within five years.

Labor is the number one cost of any IT organization. Technology components tend to decline in unit price over time, whereas labor rises over time. Thus, labor savings is a key component to any cost of ownership calculation. Figure 3 illustrates the continuing labor savings that this customer experienced over three years.

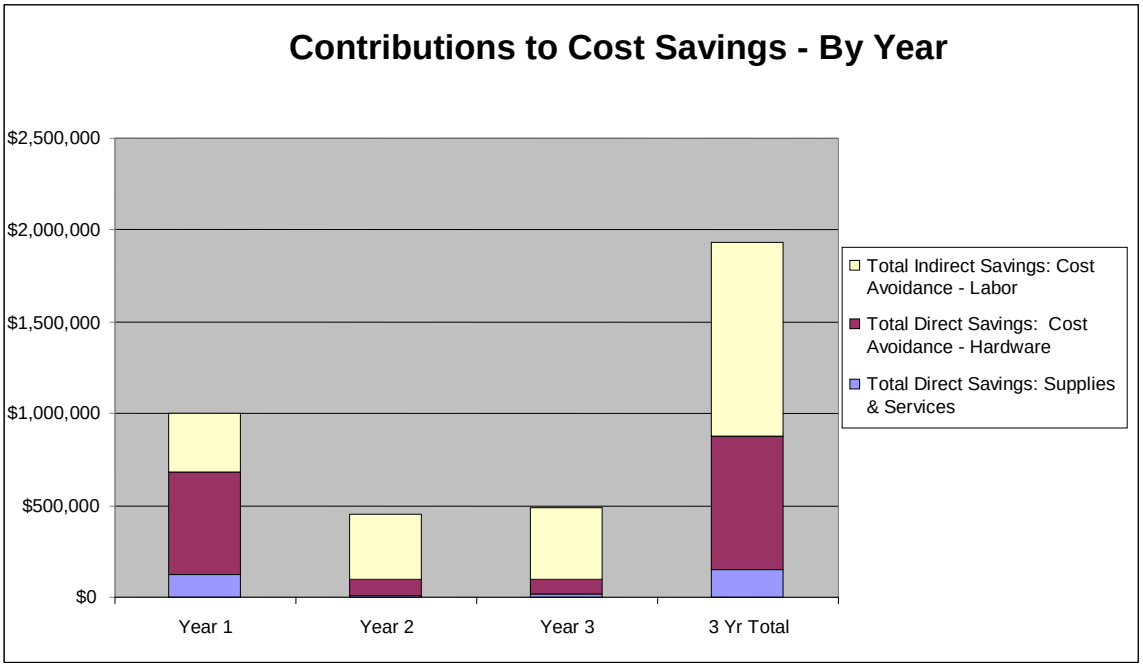

Figure 3: Contribution to savings by year for geospatial data customer

After implementing the Data Domain systems, the customer was able to reduce its managed media pool from 4,300 tapes to 2,800 by eliminating daily and weekly backups. It previously purchased 1,000 new tapes per year, whereas no new tapes are purchased now. Moreover, the organization expects to reduce the number of daily backup tapes to zero over three years, although some tape will be used for archive purposes after that time. This has resulted in further savings by eliminating the need to transfer tapes offsite on a daily basis.

Payback on the investment was the final criterion of the customer's justification. For most organizations, a payback under 12 months is clear justification. In this organization's case, the payback for the initial implementation was just three months. The IT staff expects to realize a 61% cost savings over a three-year period. Figure 4 illustrates the analysis.

Savings:	**Year 1**	**Year 2**	**Year 3**	**Total**
Direct Savings – Supplies & Services	$122,760	$13,068	$14,375	$150,203
Cost Avoidance – Hardware	560,736	85,536	85,536	731,808
Cost Avoidance – Labor	318,256	350,082	385,090	1,053,427
Total Savings	**$1,001,752**	**$448,686**	**$485,001**	**$1,935,438**
Costs:				
Total Incremental Cost of D2D	$291,000	$735,480	$175,338	$1,201,818
Summary:				
Total Net Savings of D2D	**$710,752**	**-$286,794**	**$309,663**	**$733,620**
Payback (in months)				**3**
ROI (3 years)				**61%**

Figure 4: Payback analysis for geospatial data customer

The organization's IS infrastructure manager considered the Data Domain systems a much- needed solution. In addition to the direct cost and labor savings, the organization is better able to meet its business requirements. As Figure 5 demonstrates, the customer had an astounding improvement in its recovery time objective. In this case, system restore operations were reduced from seven hours to just one and one-half hours. At the same time, the backup window was consistently achieved. The organization was able to triple its online access to backed-up data from 30 to 90 days. Finally, because static data is an ideal candidate for Data Domain's deduplication technology, the customer has seen an average of 40:1 reduction in backup data volume.

	Units	**Tape Only**	**Data Domain**	**% Improvement**	**Savings Over 3 Years**
3-Year TCO per TB	Cost per TB	$708	$413	42%	$733,620
Tape Media Cost per TB	Cost per TB per year	$2,129	$743	65%	$110,880
Offsite Storage & Transportation Costs	Cost per TB	$8	$5	33%	$6,554
Backup Window	Hours	14	12	14%	
Recovery Time Objective	Minutes	420	90	79%	
Data Kept Online	Days	30	90	300%	

Figure 5: Summary of operational improvements for geospatial data customer

Case Study #2: U.S. Financial Services Customer

Financial services organizations can be among the most demanding customers because they operate in a heavily regulated environment that leaves little leeway for error or delay. Data in these organizations must be protected reliably and with assured recovery. The stakes are always higher when dealing with other people's money.

The second organization studied had all of these challenges plus the added complication of being a mid-sized institution operating in an international environment. The company is based in the US and has five offices in North America and Europe. Each office has its own backup infrastructure, but is managed by just two administrators in the home office.

This customer's biggest problem was getting the nightly backup jobs completed. The organization's data is comprised primarily of unstructured data from MS Office applications and Exchange plus some SQL Server database data. The backup process frequently required 24 hours, which caused problems in completing all of the daily jobs reliably. High data volumes led to excessive tape handling, while hardware problems stymied operations.

To solve these problems, the customer first attempted to distribute the data onto file servers and use its backup application to back up to plain disk. The results did not make enough difference to reduce total processing time to fit in the required window. The customer then engaged Data Domain and placed a DD410 in each of four satellite offices and a DD530 in the main office. The customer could not be more pleased; on a scale of 1 to 10, the customer rated the Data Domain systems as a 10.

The factor that makes this TCO case study so compelling is that it does not rely on tape hardware purchase avoidance to offset the Data Domain purchase. Although the customer's director of technology acknowledges that he would have had to upgrade his tape infrastructure had he not purchased the Data Domain systems, he did not price out this option. Therefore, tape upgrade costs were not included in the calculations. If they had been included, the cost justification would have been even more compelling.

From a cost perspective, the greatest savings for this customer came from less tape handling labor and reduced media costs. After implementing the Data Domain systems, the customer eliminated nightly and weekly tape backups. Although tape is still used, it is used only for monthly backups and archival.

Figure 6 details these cost savings, showing the customer received about a two-year payback and a 16% three-year ROI, even without considering offsetting hardware avoidance.

Savings:	**Year 1**	**Year 2**	**Year 3**	**Total**
Direct Savings – Supplies & Services	$10,800	$3,300	$3,960	$18,060
Cost Avoidance – Hardware	0	0	0	0
Cost Avoidance – Labor	43,680	52,416	62,899	158,995
Total Savings	**$54,480**	**$55,716**	**$66,859**	**$177,055**
Costs:				
Total Incremental Cost of D2D	$112,000	$20,160	$20,160	$152,320
Summary:				
Total Net Savings of D2D	**-$57,520**	**$35,556**	**$46,699**	**$24,735**
Payback (in months)				**25**
ROI (3 years)				**16%**

Figure 6: Payback analysis for financial services customer

Figure 7 presents additional interesting results from this customer's experience, showing the customer saw a dramatic 92% reduction in tape media costs. Moreover, the backup window was reduced from 24 hours to just nine hours. Tape recoveries have been virtually eliminated. In fact, in more than a year of using the Data Domain systems, the customer has never had to perform a restore from tape. The customer has now reduced tape handing time from 10 hours per week to just two hours per month. Restore times have also been reduced by an average of 83%.

	Units	**Tape Only**	**Data Domain**	**% Improvement**	**Savings Over 3 Years**
3-Year TCO per TB	Cost per TB	$32,351	$31,125	4%	$44,610
Tape Media Cost per TB	Cost/ TB/ year	$4,375	$350	92%	$8,050
Offsite Storage & Transportation Costs	Cost per TB	$0	$0	NA	$0
Backup Window	Hours	24	9	63%	
Recovery Time Objective	Minutes/ hours	60	10	83%	

Figure 7: Summary of operational improvements for financial services customer

Although these operational improvements would have been enough for the customer to justify the purchase, labor savings provided a cost savings bonus. Figure 8 demonstrates how reduced tape handling almost single-handedly improved the customer's financial picture.

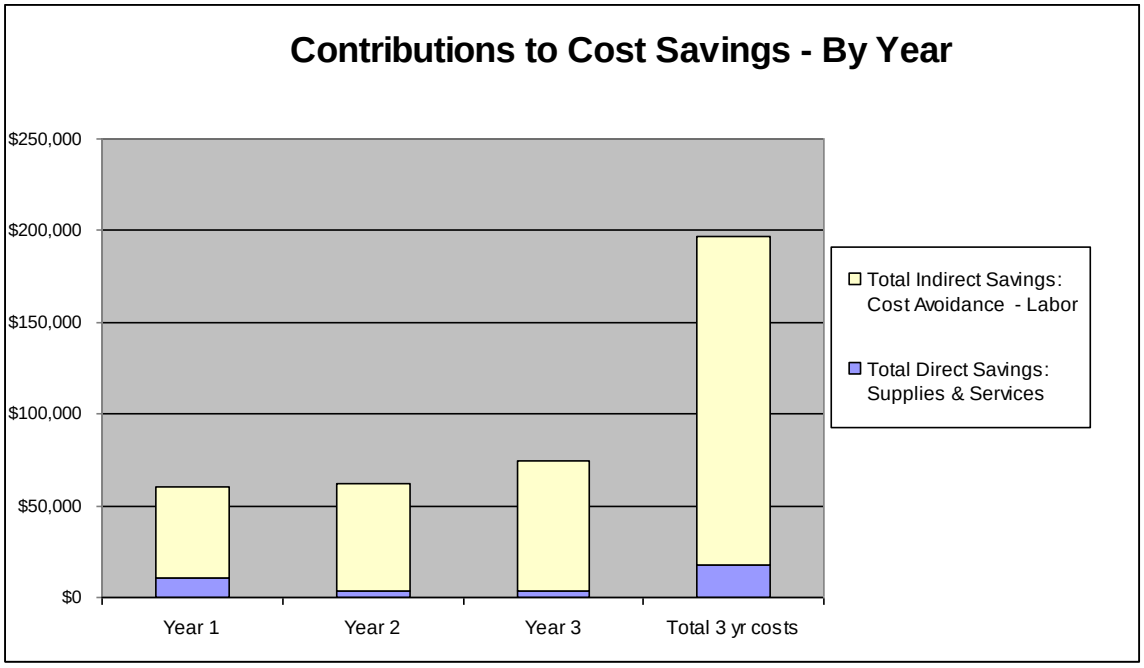

Figure 8: Contribution to savings by year for financial services customer

This labor savings has almost certainly allowed the customer to delay future head count increases related to storage management. As Figure 9 demonstrates, the savings was more than $50,000 per year, a large portion of one staff-year in salary.

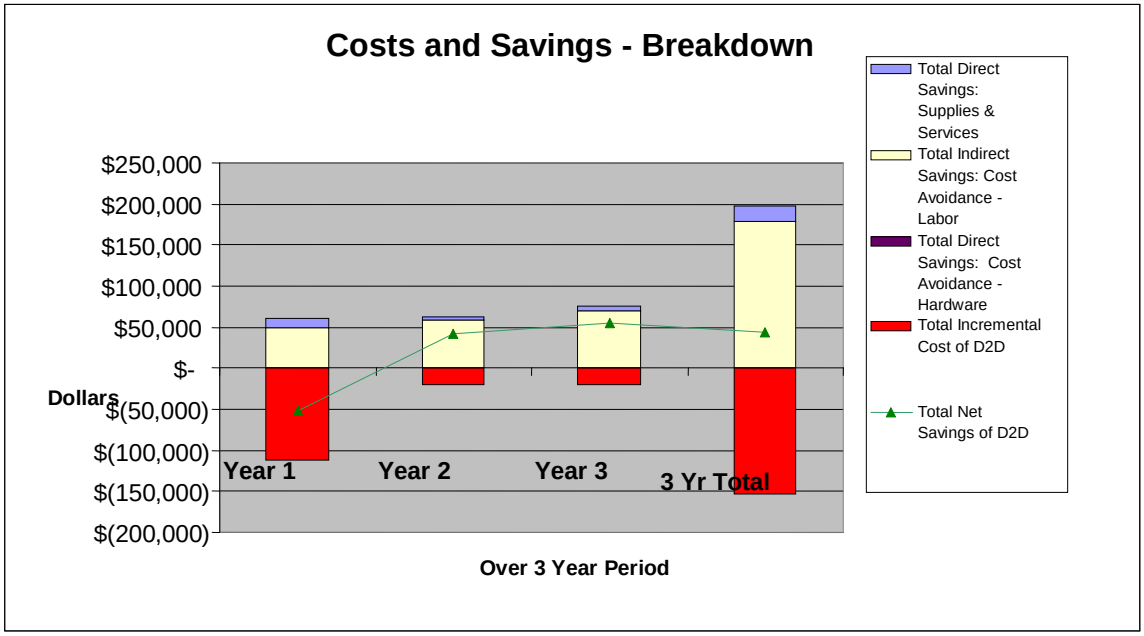

Figure 9: Breakdown of costs and savings for financial services customer

The biggest value of the Data Domain systems for this customer was the resulting backup reliability. According to the director of technology, the Data Domain systems were simple to use and "just worked." Backup performance improved in the first week after implementation.

Case Study #3: North American Oil and Gas Customer

Petroleum organizations go where the oil and gas is. This geographical necessity can lead to far-flung operations in remote locations and in climates that are inhospitable to both humans and computers due to either extreme heat or cold. Moreover, some remote locations have poor communications infrastructures and are not on the typical pick-up route for offsite tape storage and delivery companies. To top it off, the cost of labor in these locations is relatively high.

The company in the third case study faces all of these challenges. This North American company employs nearly 6,000 people and expects that number to double over the next several years. In addition to the central IT data center, this organization has seven major remote sites and eight minor remote sites in North America. The major sites have their own IT infrastructure dedicated to gathering data and sending it back to the data center.

The infrastructure of this organization is considerable. It has over 500 UNIX servers (HP-UX, Solaris, and Linux) as well as over 500 Windows servers. About 120 of these Windows servers run multiple virtual machines. The online storage arrays hold 115 TB of disk space, of which 86 TB are utilized. Each of the seven major remote locations has its own tape library for backup.

Because of the challenges associated with remote locations, in 2005 company management began exploring disk-based backup devices to improve its data protection operations. They examined various potential solutions, including one from Data Domain. In the end, no solutions other than Data Domain's were seriously considered for purchase.

The key benefits of the Data Domain solutions were threefold:

- Deduplicate data so as to minimize data sent over relatively slow communications links
- Eliminate the need to have tape devices in the remote locations
- Eliminate the need for people to manage the remote tape devices

Additional benefits realized by the company included better service level delivery and the elimination of tape media where transfer of the cartridges in extreme cold could damage them.

Once Data Domain was selected, the company initially purchased two Data Domain DD460 deduplication storage systems. As planned, after the successful initial implementation, it began a phased rollout of Data Domain systems to the remote locations starting with two DD560s in Phase 2. In Phase 3, it added three DD510s and one DD580. Both of the DD560s and the DD580 have been augmented with expansion shelves. Subsequently, one of the company's sister organizations has implemented a multi-million dollar Data Domain infrastructure to improve its SAP backup infrastructure.

Similar to the situation in the other two case studies, cost was not the primary impetus for the product purchase. Nevertheless, this customer saw a substantial savings after implementation. Here, again, the greatest savings was in labor related to the management of tape and tape devices. The company was able to eliminate the storage administrator role in all seven remote locations.

Figure 10 illustrates how much of the customer's savings was due to reduced labor. In this case, fully 94% of the savings was from labor. The customer's labor cost was $106 per hour.

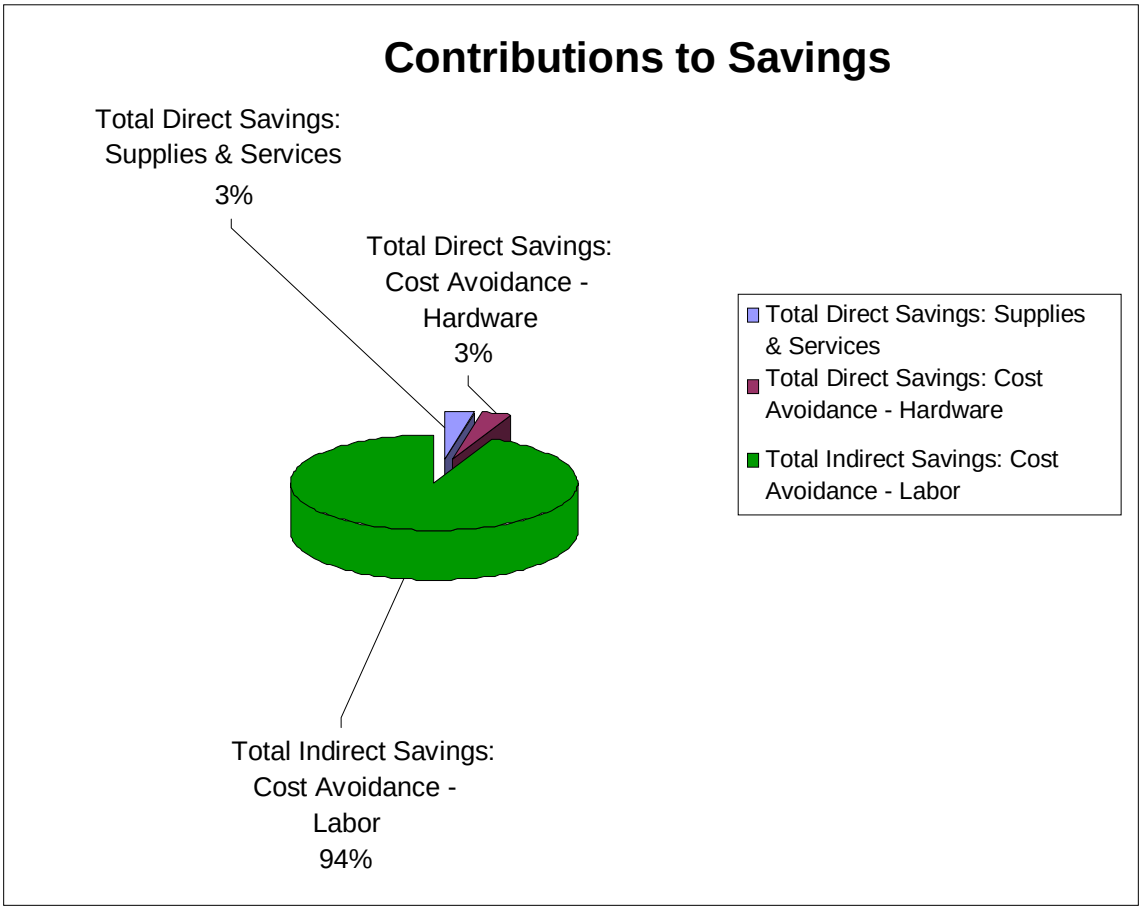

Figure 10: Contribution to savings by category for oil and gas customer

This organization's experience is instructive in other ways as well. As shown in Figures 1 and 9 for the other two case studies, the investment in Data Domain systems was very large in year 1, but offset in years 2 and 3. In this case study, the investment was more evenly spread over three years. As a result, the company saw a cost savings in every year of implementation. Figure 11 compares the cost of the systems to the realized savings.

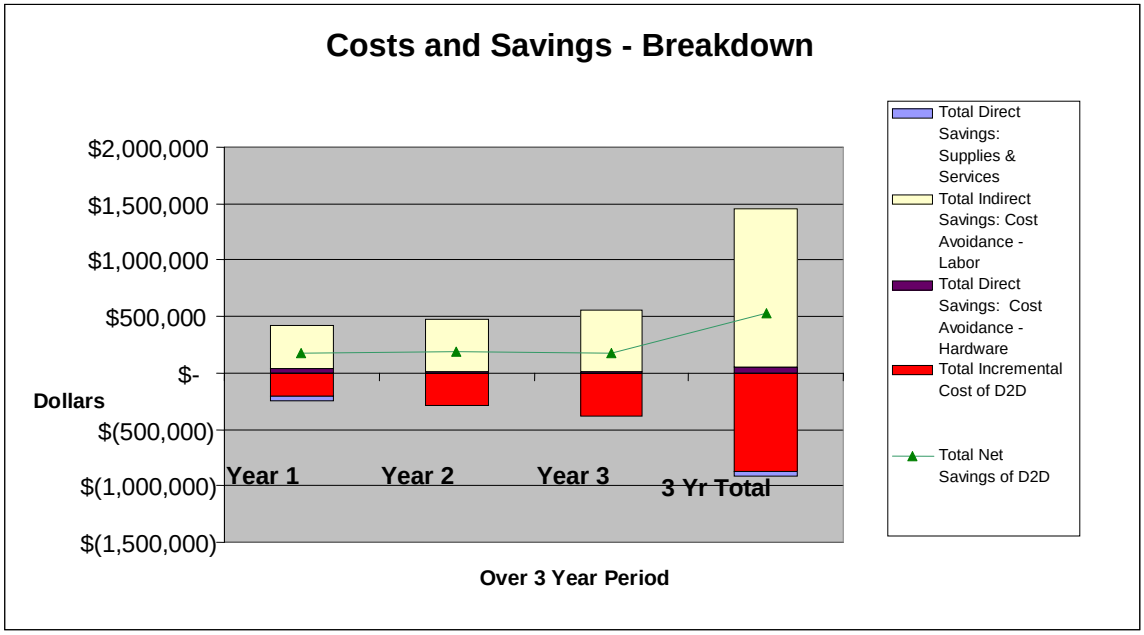

Figure 11: Breakdown of costs and savings for oil and gas customer

Even with an investment of $867,000, the customer saw a 62% ROI in the three years of ownership. Figure 12 details this ROI result. Readers should note the substantial savings from labor (largely through the elimination of remote storage administrators). Payback for the initial investment was just six months. Figures 12 and 13 show that the tape media cost increased, but that increase was due to a change in retention policy, increasing it from 13 months to 10 years due to regulatory requirements.

Savings:	**Year 1**	**Year 2**	**Year 3**	**Total**
Direct Savings – Supplies & Services	-$50,000	0	0	-$50,000
Cost Avoidance – Hardware	37,170	5,670	5,670	48,510
Cost Avoidance – Labor	385,840	463,008	555,610	1,404,458
Total Savings	**$373,010**	**$468,678**	**$561,280**	**$1,402,968**
Costs:				
Total Incremental Cost of D2D	$200,000	$286,000	$381,000	$867,000
Summary:				
Total Net Savings of D2D	**$173,010**	**$182,678**	**$180,280**	**$535,968**
Payback (in months)				**6**
ROI (3 years)				**62%**

Figure 12: Payback analysis for oil and gas customer

Readers will note that the customer's savings from tape infrastructure avoidance was relatively small, just over $48,000 (largely due to the change in data retention policy as described above, and unrelated to the implementation of Data Domain). As a result, the customer did not see savings from tape media as it had expected. It was, however, able to eliminate all tape libraries and the associated labor (i.e. storage administrators) from all seven remote locations. Figure 13 shows the three-year cost savings of more than $500,000, yielding a 434% improvement over tape alone.

	Units	Tape Only	Data Domain	% Improvement	Savings Over 3 Years
3-Year TCO per TB	Cost per TB	$295	-$985	434%	$535,968
Tape Media Cost per TB	Cost per TB per year	$872	$1,453	-67%	($50,000)
Backup Window	Hours	18	8	56%	

Figure 13: Summary of operational improvements for oil and gas customer

In addition to the substantial cost improvement, the customer's backup operations were completed in just eight hours compared to 18 hours with tape. Although the customer anecdotally believes that the Data Domains systems have improved the service level delivery and RTO, it has not empirically measured those results.

This customer has been very pleased not only with the Data Domain hardware, but with the technical support as well.

Conclusion

For regulatory and business continuity reasons, efficient, reliable data protection is an imperative for every data center. Data protection was compared to insurance in the Executive Summary. This analogy breaks down because insurance cannot contribute to a more efficient organization the way that improved backup operations can. The case studies in this white paper demonstrate that the addition of Data Domain backup appliances not only provided more reliable data protection but also actually reduced costs as well.

Direct cost savings were an important part of the business case justification. However, as these cases showed, the biggest savings was in labor. Backup managers spent less time, and in some cases dramatically less time, correcting backup failures and managing tape media. These improvements allowed the operators to shift their focus to more interesting and meaningful work while actually improving the data protection scheme. These Data Domain customers exemplify more reliable backups, more productive IT staff, and reduced costs — the data protection triple play.

Appendix A: Effects of Deduplication Storage on TCO Components

TCO Category	TCO Component	Effect of Deduplication Storage	Calculation of Costs/Savings
Hardware	Tape Backup Hardware & Maintenance	Reduction or elimination of need for any or additional tape libraries, drives, or media servers in local and/or remote offices.	No additional tape hardware, possible elimination of current hardware, and avoidance of future hardware
	Dedupe Backup Hardware & Maintenance	Incremental cost of deduplication hardware for storage and WAN vaulting/replication.	Incremental initial costs plus any additional required over 3 years
Software	Backup Software Licenses & Maintenance	With deduplication storage, no additional software. Avoids cost of additional backup licenses.	Subtract cost of additional licenses required by tape backup
Support	Labor (Backup Admin FTEs)	Reduced labor in tape mounting, handling, and transporting from remote offices.	Number of hours saved per week
	Labor (Sysadmin, Backup Admin FTEs)	Time saved due to faster restores.	Number of restores per week times number of hours saved per restore due to data being kept online
	Labor (Sysadmin, Backup Admin FTEs	End-user time saved per year due to faster restores.	Number of users affected, times Number of restores per week
Supplies	Tape Media	Reduction in number of tapes.	Reduced number of tapes in inventory and added per year (after implementing D2D) times cost of tape
Services	Offsite Tape Storage & Transportation	Reduction in storage, transportation, and tape recall costs. Potential elimination of service contracts at remote sites.	Average reduction in invoiced costs after implementing D2D

Appendix B: Case Study Details

	#1	#2	#3
Amount of Storage in TB	750	10	115
Amount of Data in TB	80	2	86
% Growth over 3 Years	10%,10%,10%	20%, 20%, 20%	20%, 20%, 20%
Number of tapes before Data Domain	4300	250	1500
Number of tapes after Data Domain	1500	20	2500*
Offsite Storage Costs/Yr before Data Domain	$5,940	Self Storage	Self Storage
Offsite Storage Reduction w/ Data Domain (Yr 1)	$1,980	Self Storage	Self Storage
# FTEs for Backup and Support	1	2	7
Tape Handing Hours Saved per Year	624	494	3640
Admin Time Saved/Yr due to faster Restores	610	208	NA
User Time Saved/Yr due to faster Restores	6560	NA	NA
Data Kept Online after Data Domain	3 months	2 months	1 month
Data Reduction Ratio with Data Domain	20:1	20:1	10:1
Backup Window Before/ After Data Domain	14 hrs to 12 hrs	24 hrs to 9 hrs	18 hrs to 8 hrs
Recovery Time Improvement	420 min to 90 min	60 min to 10 min	NA

* Regulatory requirements changed retention policy from 13 months to 10 years

About Data Domain

Data Domain® is a leading provider of deduplication storage systems. Data Domain delivers the performance, reliability and scalability to address the data protection and nearline storage needs of enterprises of all sizes. Data Domain products integrate into existing customer infrastructures and are compatible with leading enterprise backup and archive software products. To find out more about Data Domain, visit www.datadomain.com. Data Domain is headquartered at 2421 Mission College Blvd., Santa Clara, CA 95054 and can be contacted by phone at 1-866-933-3873.

Other Related Focus Research

Focus has published the following related reports. For more information, or to browse additional Focus research please go to http://www.focusonsystems.com/research.

Focus White Paper: Extending D2D to Offsite DR: The ROI Case for WAN Vaulting

About Focus

Barb Goldworm**,** president and chief analyst of Focus, has spent 30 years in technical, development, marketing, sales, senior management, and industry analyst positions with IBM, Novell, StorageTek, EMA, and multiple startups. Barb is virtualization chair for ***Interop*** and ***Blade Systems Insight***, chaired the ***2007 Server Blade Summit on Blades and Virtualization***, created and chaired the Network Storage Track of *Interop*, and has been one of the top ranked expert speakers at Data Center Decisions and SNW. Barb has been a regular expert columnist and speaker for *TechTarget, Ziff-Davis, Computerworld Storage Networking World Online, Network World* and *Virtual Strategy Magazine*. Co-author of ***Blade Servers and Virtualization: Transforming Enterprise Computing While Cutting Costs***, she has published extensively, including research reports, market studies, landscape reports, and white papers.

Phil Goodwin, contributing analyst, has over 25 years of industry experience and specializes in data protection and storage management best practices.

Focus delivers research, analysis and consulting, focused on systems, software and storage in the enterprise, SMB, and government markets. Focus areas include: Systems, Storage and Enterprise Management (Physical and Virtual); Server, Desktop and Application Virtualization/Streaming; High Availability, Disaster Recovery, Business Continuity, Backup, Data Protection; Storage Networking (NAS, SAN, Fibre Channel, iSCSI); Storage, Networks, I/O and File Virtualization; Storage Technologies (Clustered File Systems, Data deduplication, VTL, etc.); Blade Systems (Server, Workstation and PC Blades); and Business Benefits of Technology (ROI, TCO). **www.focusonsystems.com**