

Lab Validation Report

HP Data Protector and Deduplication Solutions

Scalability and Performance from the Core to the Edge

By Tony Palmer

June 2010

Contents

Introduction.....	3
Background	3
HP Data Protection Solutions.....	4
ESG Lab Validation.....	6
Ease of Deployment and Integration	6
Scalability and Performance	9
Remote Office Protection	11
Cost-Efficient Protection	16
ESG Lab Validation Highlights.....	18
Issues to Consider.....	18
The Bigger Truth	19
Appendix.....	20

ESG Lab Reports

The goal of ESG Lab reports is to educate IT professionals about emerging technologies and products in the storage, data management and information security industries. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments. This ESG Lab report was sponsored by Hewlett Packard.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. Copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482.0188.

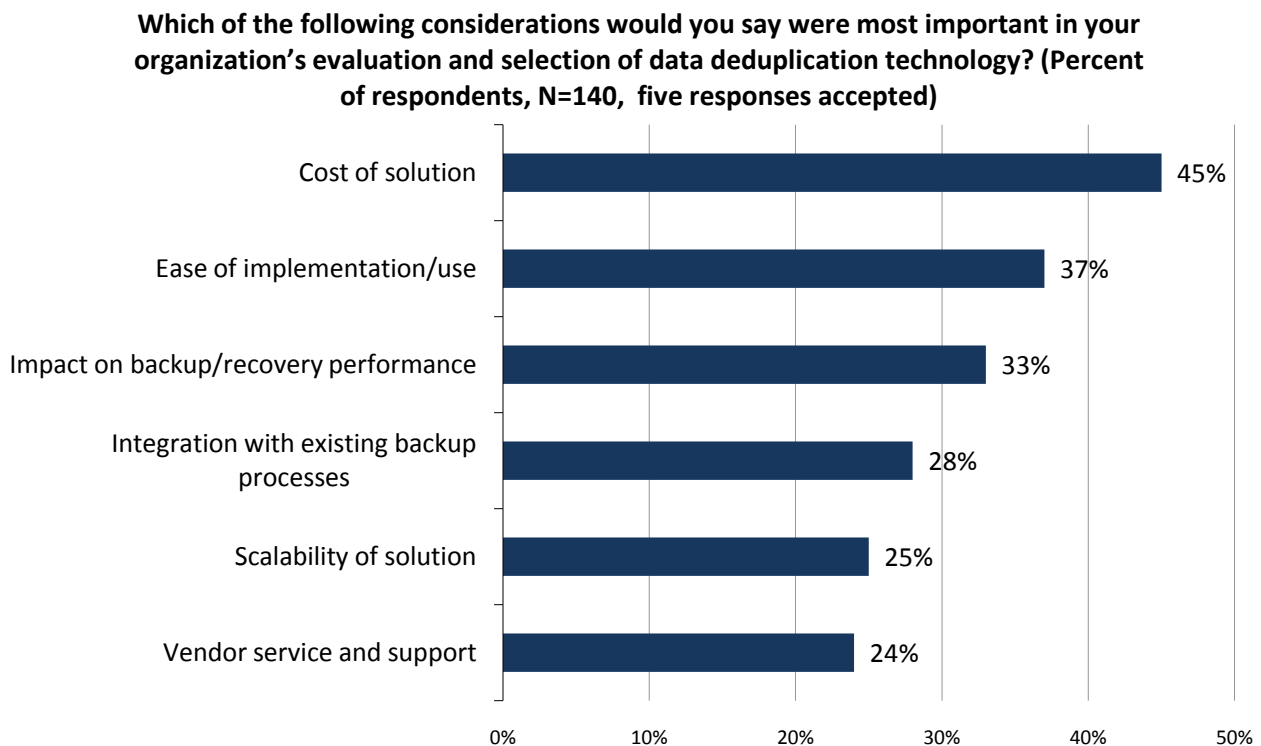
Introduction

A large number of organizations have deployed disk-to-disk backup technologies to improve the speed and reliability of their backup and disaster recovery operations. A growing number of these organizations look to data deduplication to enhance retention periods and reduce the cost of storage for backups and disaster recovery. This ESG Lab Validation Report examines Hewlett Packard's family of backup and recovery solutions that combine the power of HP StorageWorks Virtual Library Systems (VLS) in the data center and the agility of HP D2D appliances in remote offices, tied together with HP Data Protector backup and recovery software. Special attention was paid to ease of implementation as well as the solution's ability to improve the speed and reliability of disk-based data protection while reducing the cost of disk capacity and network bandwidth. Some of the issues associated with choosing a deduplication solution are also explored.

Background

While deduplication can reduce the cost of the raw storage required to store and replicate backup data on disk, integration with the existing ecosystem is crucial. As shown in Figure 1, recently completed ESG research indicates that ease of implementation, performance impacts, and integration with existing backup processes are key concerns.¹ Robust management, edge to core replication, tape integration, and deduplication options are important considerations as well, especially within large enterprise-class organizations. The diverse family of Hewlett Packard data protection solutions is ideally suited to address these, and other, concerns.

Figure 1. Data Deduplication Evaluation Criteria



Source: Enterprise Strategy Group, 2010.

Deduplication

Choosing a deduplication strategy should include a discussion on how and where deduplication should occur. There are two basic locations deduplication can occur: at the source, typically accomplished via a software agent running on the client machine, or at the target, which involves either writing directly to the device or running a

¹ Source: ESG Research Report, *Data Protection Survey*, to be published in Q2 2010.

software agent on the media server to perform deduplication. All deduplication includes some level of overhead. If it occurs at the source, that overhead occurs on the client machine or media server and may have an impact on backup performance due to the software required on client systems, which can consume processing, storage, and/or network resources to deduplicate data. When deduplication is performed at the target, the overhead is incurred in the device where data is being written.

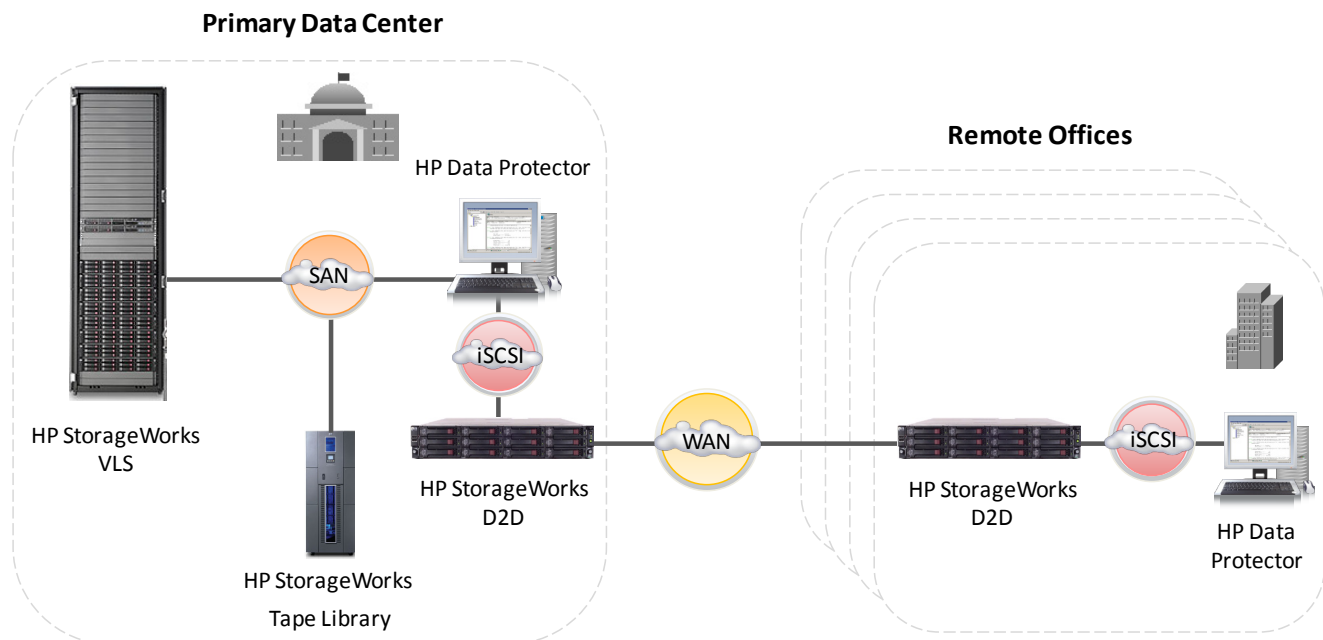
In addition, it is important to differentiate between a pure software approach — in which the software runs on an industry standard platform, — and an appliance-based (or storage hardware-based) solution. With a software solution users have flexibility in their choice of physical storage, but must ensure that the system has enough I/O bandwidth to support their performance needs, sufficient system resources to support desired deduplication rates, and the right storage security to prevent potential data loss. An appliance-based solution has the ability to address many of these concerns, but locks users into a particular hardware platform.

HP's D2D and VLS are target-side deduplication appliances which are designed to provide cost-effective, easy-to-deploy deduplication and are tuned for optimal performance on each hardware platform.

HP Data Protection Solutions

Hewlett Packard offers diverse data protection hardware and software to address the concerns of enterprises from the core data center to the smallest remote office. Figure 2 shows the HP family of data protection solutions as they might be deployed in a typical distributed enterprise.

Figure 2. HP Data Protection Solutions



HP StorageWorks VLS virtual tape libraries provide a high performance primary backup target with optional deduplication in the FC SAN enterprise data center while HP StorageWorks D2D serves as an easy to manage data protection appliance in mid market data centers and remote offices with deduplication and WAN-efficient replication across sites. HP Data Protector software provides a solution to manage the D2D and the VLS not only in single-site backup environments but also in replicated solutions. While this report focuses on the benefits of a total HP solution, it's important to note that HP products integrate well with third party products. Data Protector works with any inline de-duplication solution and the D2D and VLS can be used with third party backup software. However the end to end, single vendor solution including replication enablement and integration is a unique HP offering.

Benefits of HP StorageWorks Data Protection Solutions

VLS:

- **Highly scalable capacity and performance:** Up to 4800 MB/sec of throughput and 1280 TB of usable storage.
- **The entire capacity of the system can be presented as a single virtual library target:** The system can scale without time consuming reconfiguration and rebalancing of backup software and backup jobs.
- **Accelerated deduplication:** Fast hardware compression, in combination with post process deduplication, provides capacity efficiency without impact to backup windows.
- **WAN efficient replication:** Fast, cost-effective disaster recovery capability between data centers and remote sites.

D2D:

- **Distributed capacity and performance:** The D2D Backup Systems offer high performance multi-streaming backup speeds of up to 720 GB/hour.
- **HP Dynamic deduplication:** Enables longer term data retention on disk and WAN efficient replication.
- **WAN efficient replication:** Fast, cost-effective disaster recovery capability between multiple remote sites.
- **Ease of use and deployment:** HP's D2D Backup Systems are designed for easy installation and deployment for mid-sized business environments—it is ready to deploy right out of the box.

As of the publication of this report, HP has refreshed the entire D2D product family and introduced a new high-end model, the D2D 4312. The D2D 4312 has more processing power and offers higher capacity than previous generation D2D models. The new D2D product family runs a new 64bit data deduplication technology called HP StoreOnce. HP's vision is to port StoreOnce deduplication technology to several HP platforms – including HP Data Protector. While ESG tested the previous generation D2D appliances, the scenarios depicted and conclusions drawn in this report still apply.

Data Protector:

- **Advanced Backup to Disk** - 24/7 information access and quick disaster recovery.
- **Multiple Recovery Point/Recovery Time Objectives** - Achieve business-driven recovery objectives.
- **Manage data effectively** - within existing budgets and infrastructure, even as the quantity of data grows.
- **Centralized Data Protection** - Protect data on distributed physical and virtual infrastructures.
- **Broad Interoperability** - Integrates with partner solutions including NetApp, Data Domain, IBM ProtecTIER, and supports any third party inline deduplication target appliance.

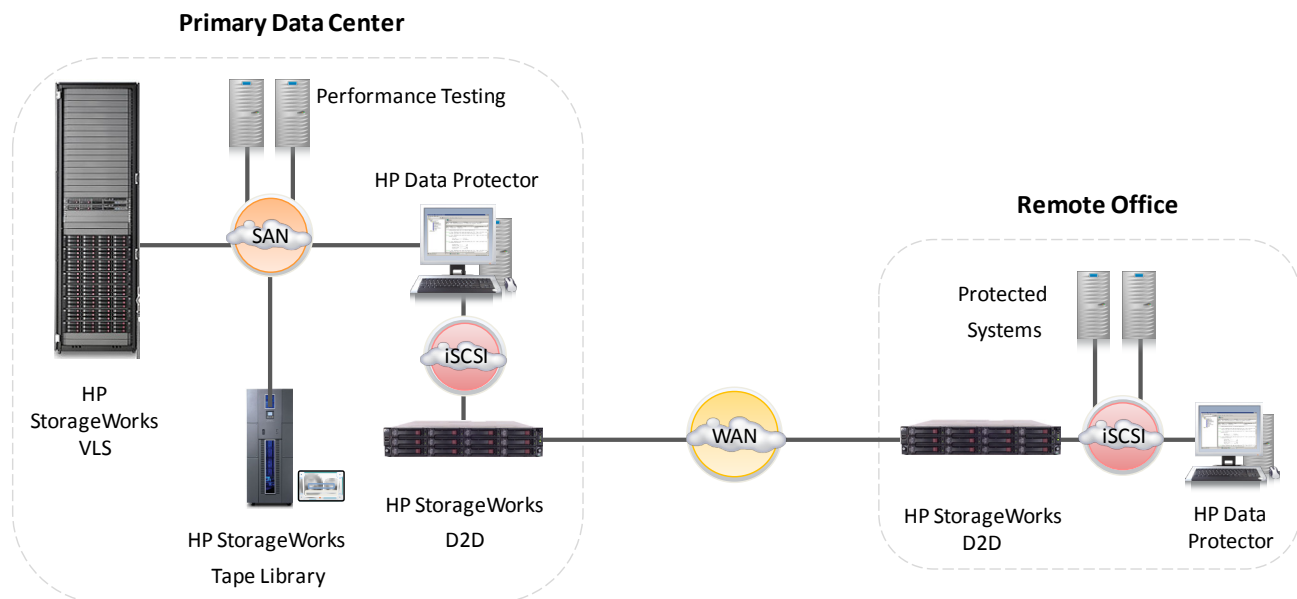
ESG Lab Validation

ESG Lab performed hands-on evaluation and testing of HP's data protection solutions at an HP facility in Fort Collins, CO. Testing was designed to demonstrate the scalability, performance, and ease of management of HP's solutions from the point of view of a typical enterprise systems administrator integrating HP's disk-based solutions into an existing tape environment.

Ease of Deployment and Integration

The test environment, shown in Figure 3, was used throughout testing. Testing began with a one-node HP StorageWorks VLS9000 array with deduplication, physically installed and powered up as it would be by HP professional services for an enterprise customer, in a data center environment with HP Data Protector software installed. Other elements typical of an enterprise environment, such as physical tape libraries and HP StorageWorks D2D appliances, were also present in the test bed.²

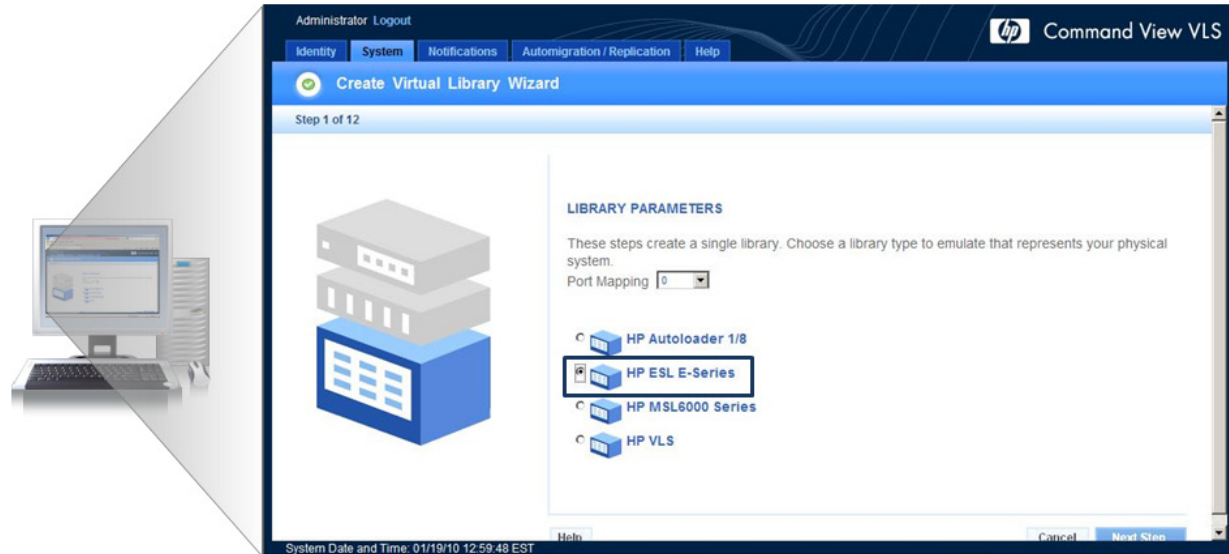
Figure 3. The ESG Lab Test Bed



ESG Lab Testing

ESG Lab initially logged in to the Command View VLS console by pointing Internet Explorer at the administrative IP address on the VLS9000 library and entering the administrator username and password. Next, the Create Virtual Library Wizard was launched. The Create Virtual Library Wizard walks the administrator through the steps to create a virtual tape library, asking for such details as library type, and is shown in Figure 4.

² Detailed configuration information can be found in the Appendix.

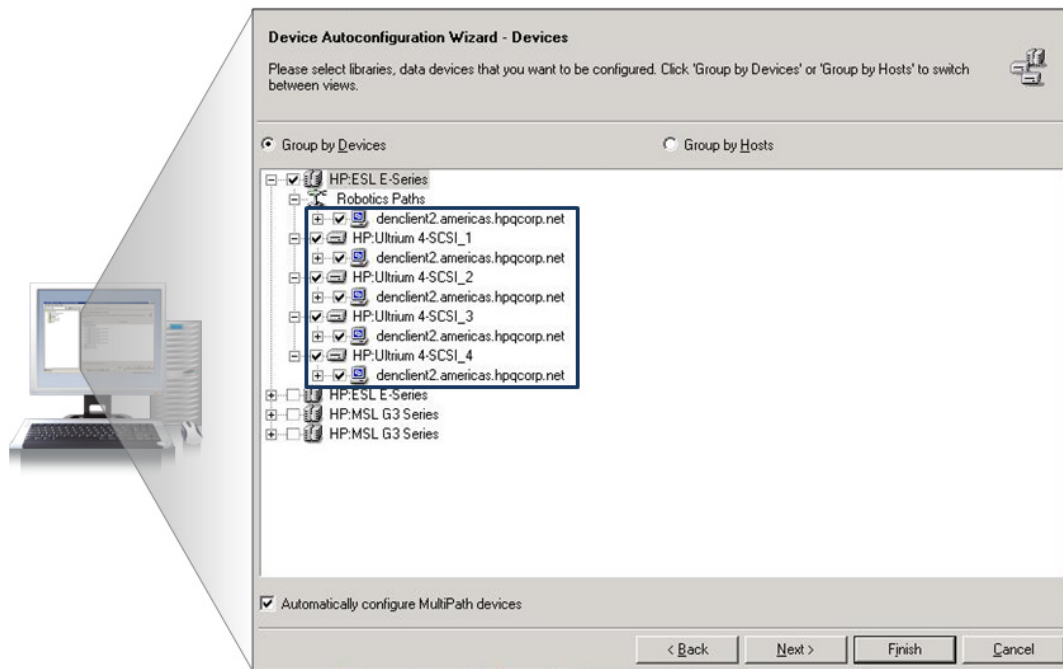
Figure 4. Creating a Virtual Tape Library with VLS

ESG Lab selected an HP ESL E-Series to match the physical library type already installed in the test environment. Next, and more important, the virtual tape drive type and quantity, as well as the capacity and quantity of virtual tape cartridges, were set, as seen in Figure 5.

Figure 5. Virtual Tape Library Created

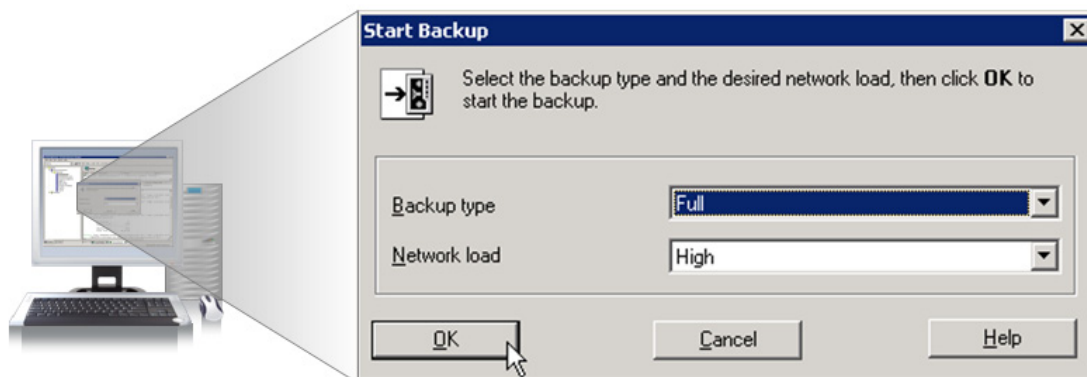
In less than two minutes, the Create Virtual Library Wizard was completed and the virtual tape library was configured and presented on the SAN. Next, ESG Lab started the HP Data Protector Autoconfigure Wizard. The HP Data Protector Autoconfigure Wizard, as seen in Figure 6, discovers new backup target devices and prepares Data Protector to use them.

Figure 6. Configuring a Virtual Tape Library in Data Protector



The Data Protector Autoconfigure Wizard took about two minutes to discover and add the Virtual Tape Library and its virtual tape drives to Data Protector. As Figure 7 shows, within five minutes of sitting down at the keyboard, ESG Lab was running a full backup of the first server to the HP VLS9000.

Figure 7. Running the First Backup In Data Protector



Why This Matters

ESG Research found that 46% of early adopters of deduplication solutions indicated ease of deployment as the single most important factor in purchasing a deduplication solution.³ This is especially important for enterprise-class appliances deployed in large, complex environments where backup policies span hundreds of servers and dozens of applications—stretching resources to the limit. ESG Lab has confirmed that an HP StorageWorks VLS is extremely easy to configure and manage. The system was dropped into an existing tape environment and performing backups in less than five minutes using familiar tools and methodologies.

³ Source: ESG Research Report, *Data Protection Survey*, to be published in Q2 2010.

Scalability and Performance

One of the fundamental advantages of VTL backup is the ability to run many backup streams concurrently using multiple virtual tape drives. A single tape drive can only perform one backup at a time. To get more than one backup job running at the same time, more tape drives must be added and run in parallel. A disk-based backup and recovery solution with many random access disk drives emulating many virtual tape drives can run many backup jobs simultaneously. The random access nature of disk also provides improved performance when locating individual files to be restored.

ESG Lab Testing

ESG Lab performed backups using a single node VLS9000 as a target first and then repeated the test after upgrading the VLS to two nodes in order to examine its relative performance as storage capacity is scaled. A full backup of multiple servers was simulated using the HP tapeperf utility set to generate data with 2:1 compressibility.⁴ The first iteration ran with two servers running ten backup streams; the second iteration was performed with three servers running sixteen streams. Performance scaled linearly when the second node was added. The screen capture in Figure 8 shows the VLS Command View Console during the two node test, running at 1,164 MB/sec.

Figure 8. Two Node VLS System Performance

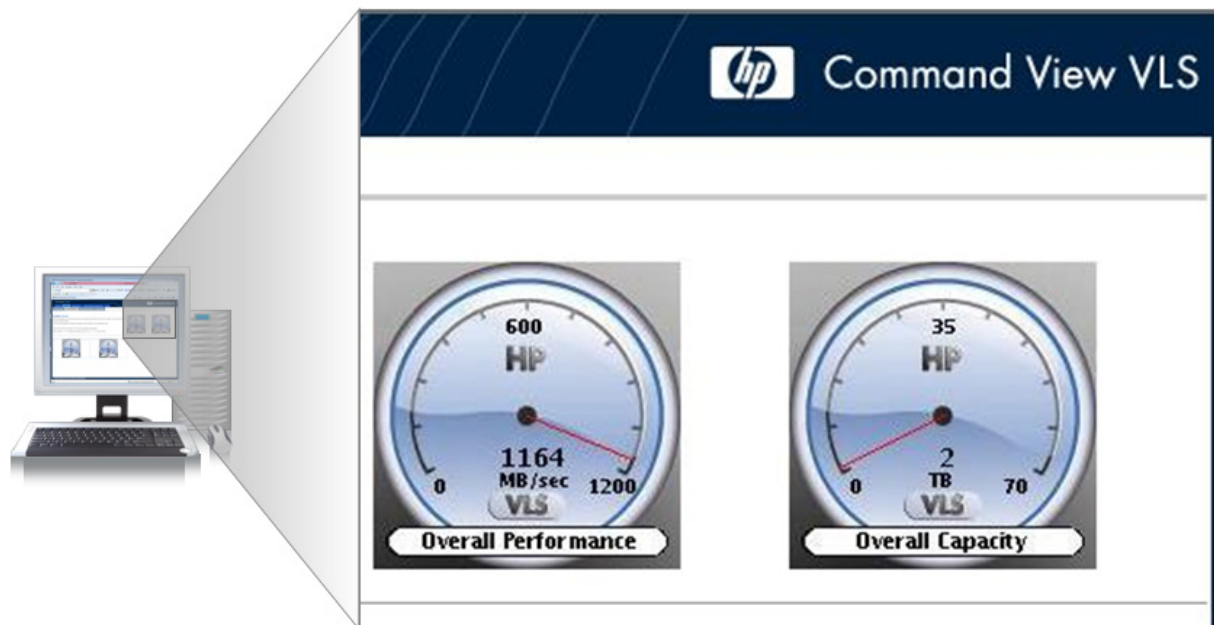


Figure 9 shows actual performance results obtained in ESG Lab testing projected out to a fully populated, eight-node system. Detailed results for each test run are shown in Table 1.

⁴ Full testing configuration is described in detail in the Appendix.

Figure 9. VLS Performance Scaling

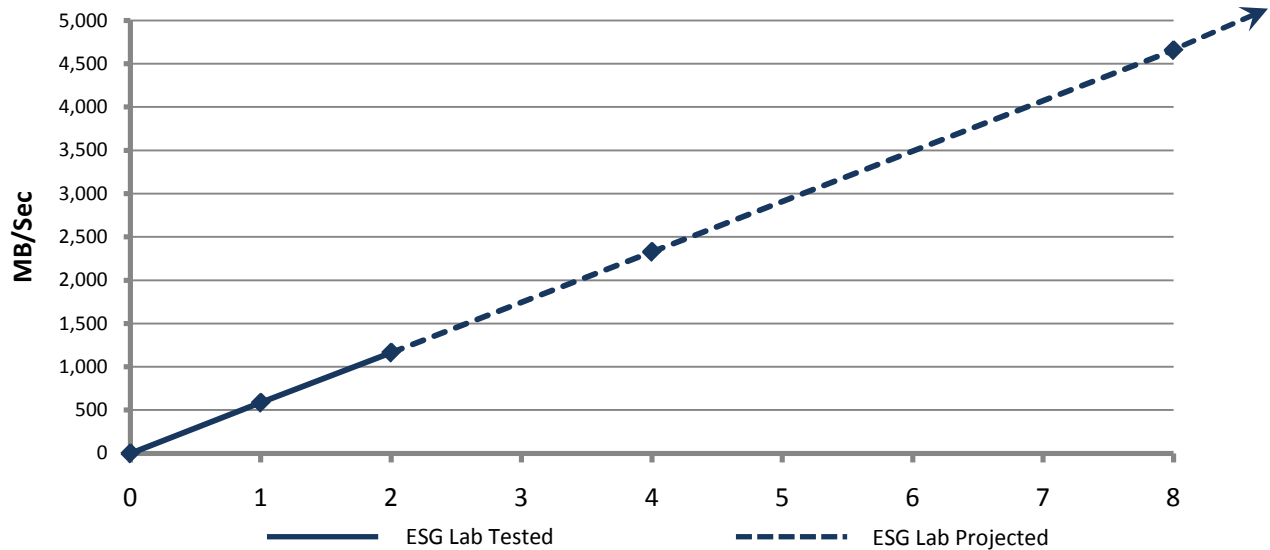


Table 1: Raw VLS9000 Backup Results

Active Backup Streams	Active Virtual Tape Drives	Throughput (MB/sec)
10	10	585
16	16	1,164

What the Numbers Mean

- When a node was added to the VLS, performance scaled nearly linearly with no degradation.
- An eight-node system should be able to achieve 4,656 MB/sec of sustained backup throughput, representing the ability to protect nearly 128 TB of data in an eight-hour backup window.

Why This Matters

ESG research⁵ has found that the number one challenge enterprises report with their data protection processes and technologies is the need to reduce backup and recovery times. Backup administrators have been struggling for years to get nightly backups completed before business resumes in the morning. Quicker recoveries are also needed to increase user productivity and meet service level agreements.

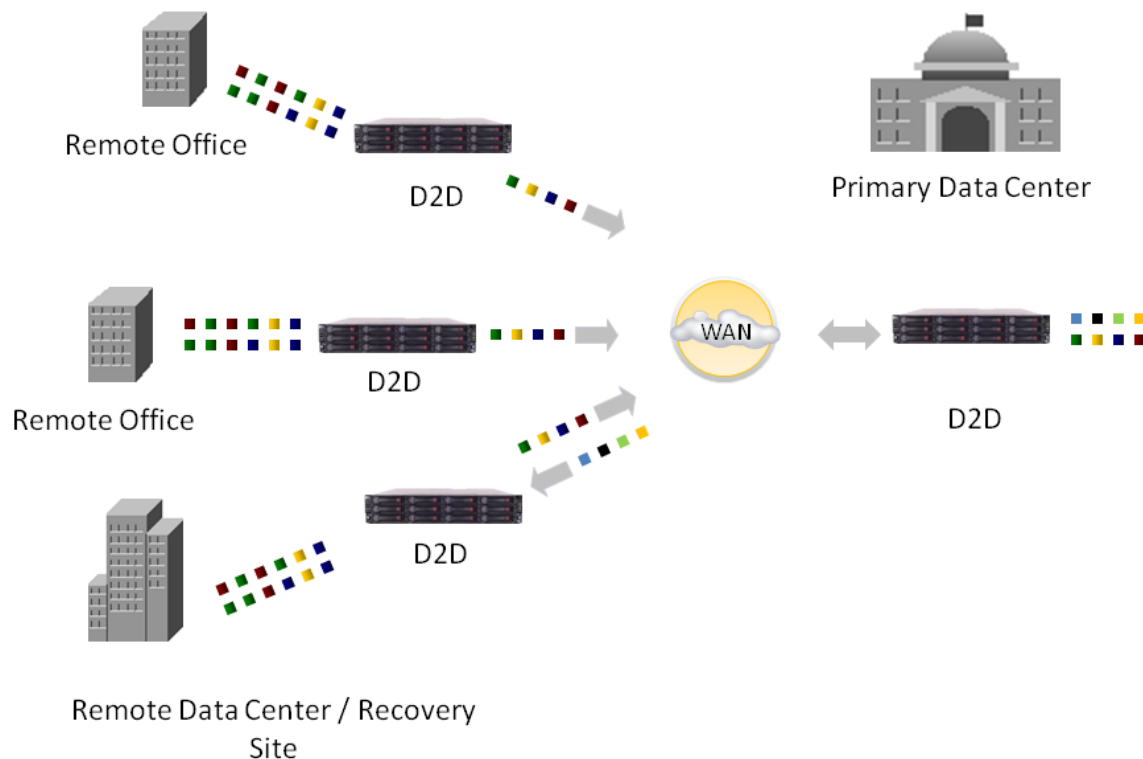
ESG Lab validated through direct test and audit that HP's VLS 9000 can linearly scale aggregate backup throughput as nodes are added. In other words, a single VLS9000 disk backup system can be used to protect up to 128 TB of data in an eight hour shift and restore individual files in a matter of seconds. Accelerated deduplication in the VLS series means that users can meet the protection requirements of a large number of servers with one system, enabling zero-impact deduplication while lowering acquisition costs and operational complexity.

⁵ Source: ESG Research Report, *Data Protection Survey*, to be published in Q2 2010.

Remote Office Protection

The family of HP disk and tape data protection solutions, combined with HP Data Protector software, can be configured to create an automated, edge-to-core data protection topology (see Figure 10) that spans multiple sites and provides disk-to-disk-to-tape (D2D2T) functionality. HP StorageWorks D2D appliances provide WAN-efficient movement of data between sites and storage tiers while Data Protector provides the single point of management and catalog for backup data—regardless of where it resides (remote office or corporate data center), what type of media it is stored on (disk or tape), or its age (recent backup or long term archive). D2D disk-based backup and replication appliances support data deduplication to reduce the resources required to store backup images on disk and replicate backup images over a WAN.

Figure 10. D2D WAN Efficient Replication



ESG Lab Testing

ESG Lab used HP Data Protector software to configure, automate, and track the migration of backup data residing on HP D2D data protection appliances, as shown previously in Figure 3. An edge-to-core D2D2T data protection strategy was implemented using an HP D2D appliance located in a simulated remote office. Remote office backup data was replicated over a simulated WAN to an HP D2D in a corporate data center with data movement carried out by the D2D systems. The object copy capabilities of the Data Protector software were used to write a copy of the data to removable media in a Fibre Channel SAN-attached HP StorageWorks MSL tape library. First, ESG Lab logged in to the D2D web management console, shown in Figure 11.

Figure 11. The D2D Management Interface

HP D2D Backup System

Home Virtual Tape Devices Configuration Status Tape Attach Replication Administration


D2DBS : IPv4 10.5.5.4 : IPv6 fe80::21e:bff:fec1:d22e/64 ✓ Status

Summary

System Information		Status	
IP Address (1)	10.5.5.4	System State	✓ OK
MAC Address (1)	00:1E:0B:C1:D2:2E	Network	✓ Network Port 1: 1Gb/s
IPv6 Address (1)	fe80::21e:bff:fec1:d22e/64		✓ Network Port 2: Not Configured
Name	D2DBS	Fibre Channel	⚠ FC Port 1: Not Used
Network Name	D2DBS.DP.local		⚠ FC Port 2: Not Used
Up Time	17 Days 17 Hours 49 Minutes	Disk Space Used	✓ 5.0 GB of 18.0 TB (0%)
Serial Number	CZC8171K17	Deduplication Ratio	✓ (10.8 : 1)
Firmware Revision	101.041	User Data Stored	✓ 53.8 GB
		Storage	✓ OK
		RAID	✓ OK
		RAID Battery	✓ OK
		Devices	✓ OK
		Tape Attach	✓ No Physical Tape Drive(s)
		Replication	✓ OK

The first backup was 'seeded' or replicated locally over a gigabit ethernet LAN connection. Seeding is often employed by users who have large data sets at remote offices as it allows the first bulk transfer of data to complete very quickly. The D2D device is then shipped to the central data center and from then on, updates require much less bandwidth thanks to deduplication. The first bulk replication is illustrated in Figure 12.


Figure 12. Initial Replication of a Remote Office Dataset



Event Details	
Message	Data job for a cartridge has completed successfully
Reason	-
Source Appliance Name	D2DBS
Source Appliance Address	---
Source Appliance Serial Number	CZC8171K17
Target Appliance Name	D2DBS-CZC9062Q6J
Target Appliance Address	10.5.5.3
Target Appliance Serial Number	CZC9062Q6J
Connection Failure	-
Mapping Name	Slot Mapping 1
Direction	Replication
Job Started	23:07 2010/01/20
Job Stopped	23:18 2010/01/20
Stage	Finished
Data Replicated	53 GB
Duration	00:10:40
Throughput	84 MB / s
Bandwidth Saving	90 %
Status Code	-

Replication of the first backup transferred 53 GB of data in 10 minutes and 40 seconds over an unrestricted Gigabit Ethernet connection. Once the first full backup was completely replicated to the target D2D appliance, ESG Lab inserted the 'Network Nightmare' WAN simulator and restricted throughput to 2 megabits per second to simulate a nearby WAN connection between the remote office and data center.

Figure 13. Capacity Efficient Replication

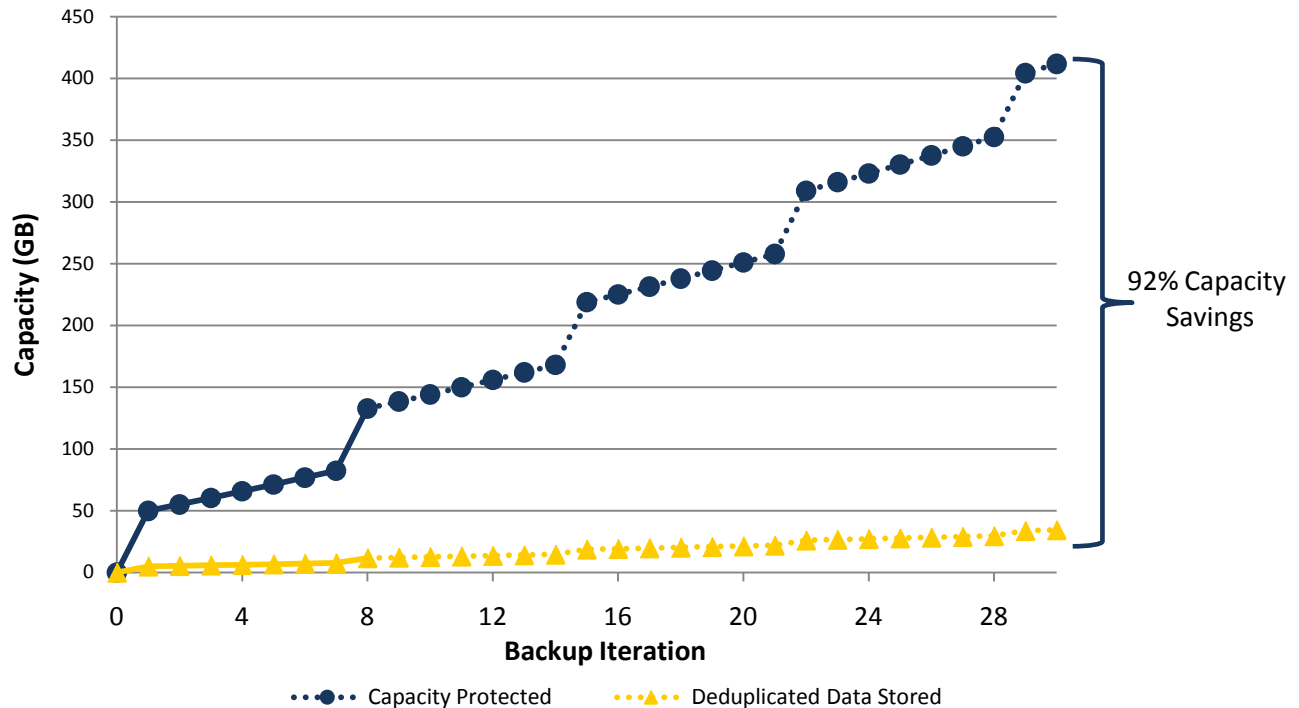


Event Details	
Message	Data job for a cartridge has completed successfully
Reason	-
Source Appliance Name	D2DBS
Source Appliance Address	---
Source Appliance Serial Number	CZC8171K17
Target Appliance Name	D2DBS-CZC9062Q6J
Target Appliance Address	10.5.5.3
Target Appliance Serial Number	CZC9062Q6J
Connection Failure	-
Mapping Name	Slot Mapping 1
Direction	Replication
Job Started	23:42 2010/01/20
Job Stopped	1:00 2010/01/21
Stage	Finished
Data Replicated	5 GB
Duration	01:17:49
Throughput	1 MB / s
Bandwidth Saving	81 %
Status Code	-

The first incremental backup was performed and automatically replicated over the 2 Mbit/sec simulated WAN connection. Replication of the incremental backup transferred 5 GB of deduplicated data in 1 hour, 17 minutes and 49 seconds. This means the D2D transferred at the equivalent of 11 Mbit/sec over the 2 Mbit/sec WAN connection. By comparison, the second full backup resulted in the equivalent of 22.9 Mbit/sec over the same link. The higher virtual throughput is due to the greater volume of duplicate data in a full backup.

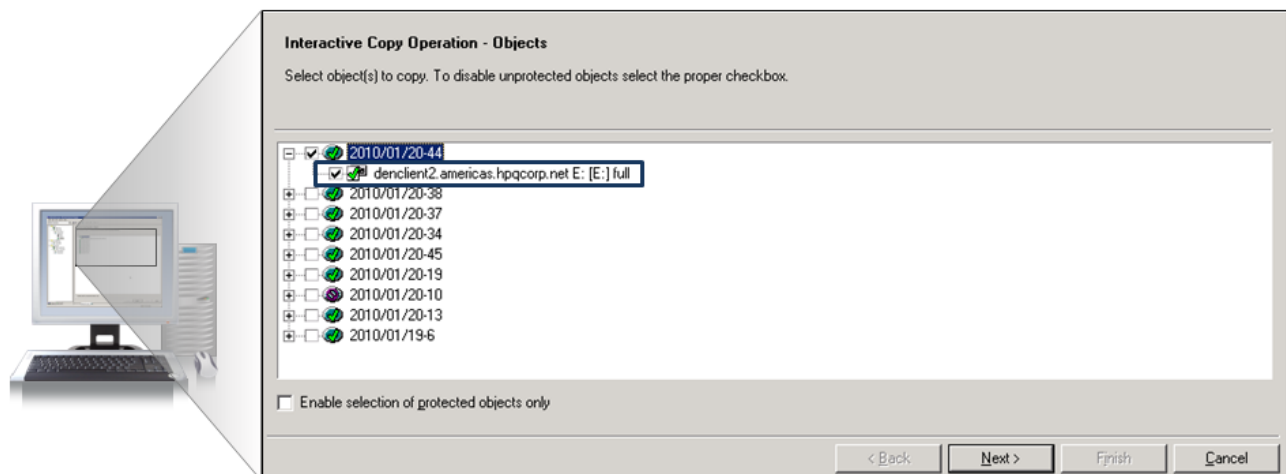
Figure 14 shows actual and projected deduplication capacity savings over 30 days of backups on a weekly full, daily incremental backup schedule. The capacity savings over 30 days was projected at 92%—an 11.94:1 deduplication ratio.

Figure 14. Deduplication Capacity Savings Over Time



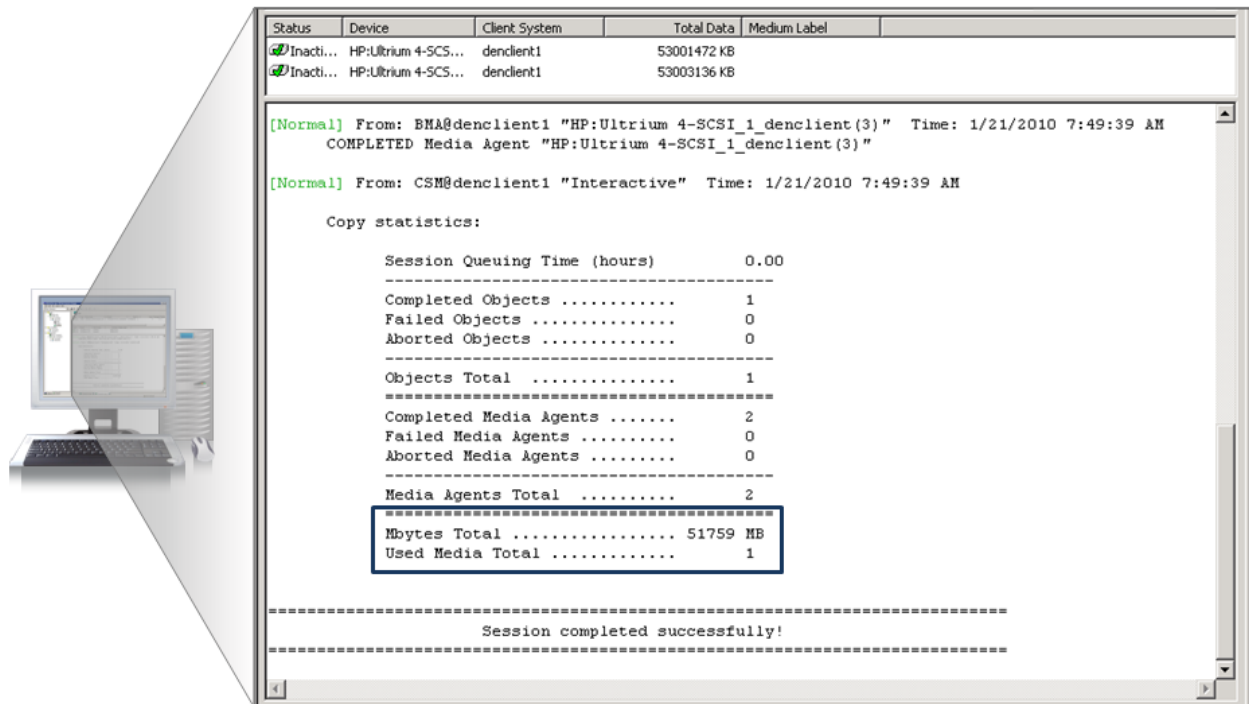
Finally, ESG Lab copied the latest full backup to tape the copy to tape capabilities of Data Protector with the D2D, which can be used for archiving. D2D frees up space by expiring source media (virtual tapes) rather than deleting the source data (files). This ensures that data that exists in multiple backups is not deleted until all references to it are deleted.

Figure 15. Copy to Tape



In 20 minutes, the copy to tape was complete.

Figure 16. Copy to Tape Complete



Why This Matters

A growing number of organizations are struggling to protect information assets residing in remote and branch offices. Most are alarmed at the rate of data growth in these locations. Many lack the IT staff and expertise needed to manage traditional tape-based protection methods. Many more are frustrated with the cost and complexity of managing tape media at remote offices. Disk-based backup and replication to a corporate data center reduces the complexity and risk, but, until recently, it's been too expensive to justify due to the cost of remote office disk capacity and the WAN connectivity required.

At the time of testing, the street price started at \$4,499 for a 2 TB system. This included dynamic deduplication and low bandwidth replication—ESG Lab found the HP Storageworks D2D to be an affordable, simple, and effective solution for the protection of valuable information assets residing in remote and branch offices. Data protector provides a single point of management and catalog for local and replicated backup data that reduces complexity and cost for distributed environments.

Cost-Efficient Protection

Organizations of all sizes are struggling to meet the conflicting challenges associated with macro-level global financial uncertainty and micro-level information storage growth and complexity. A growing number of IT managers are turning to virtualization and consolidation technologies to meet these challenges. With a focus on scalability, automated management using rich software tools, and capacity-efficient pricing, HP's data protection solutions are an excellent example of solutions that are purpose-built to address these issues.

ESG Lab created a total cost of ownership (TCO) model based on a hypothetical backup environment with multiple remote offices, a major data center, and remote replication for disaster protection. The scenario examines cost-savings associated with moving from a tape-based to a VLS and D2D-based backup and recovery strategy, although it still assumes use of tape for long-term archive of backups. Costs were broken down by category: capital expenditures, administrative costs, tape costs, maintenance costs, power and cooling costs, and total floorspace costs. The cumulative costs for both tape- and disk-based backups were calculated annually over a five year period. A number of assumptions were made and included in the calculations based on what a current IT organization might have in place for equipment, WAN connectivity, backup and restore policies, and capacity and performance requirements.⁶ Comparisons were made between the total cost of ownership of a traditional tape infrastructure with no replication or deduplication and an HP Data Protection environment with disk-based backup, deduplication, and replication.

Figure 17. The HP Data Protection Solutions Advantage Over Five Years

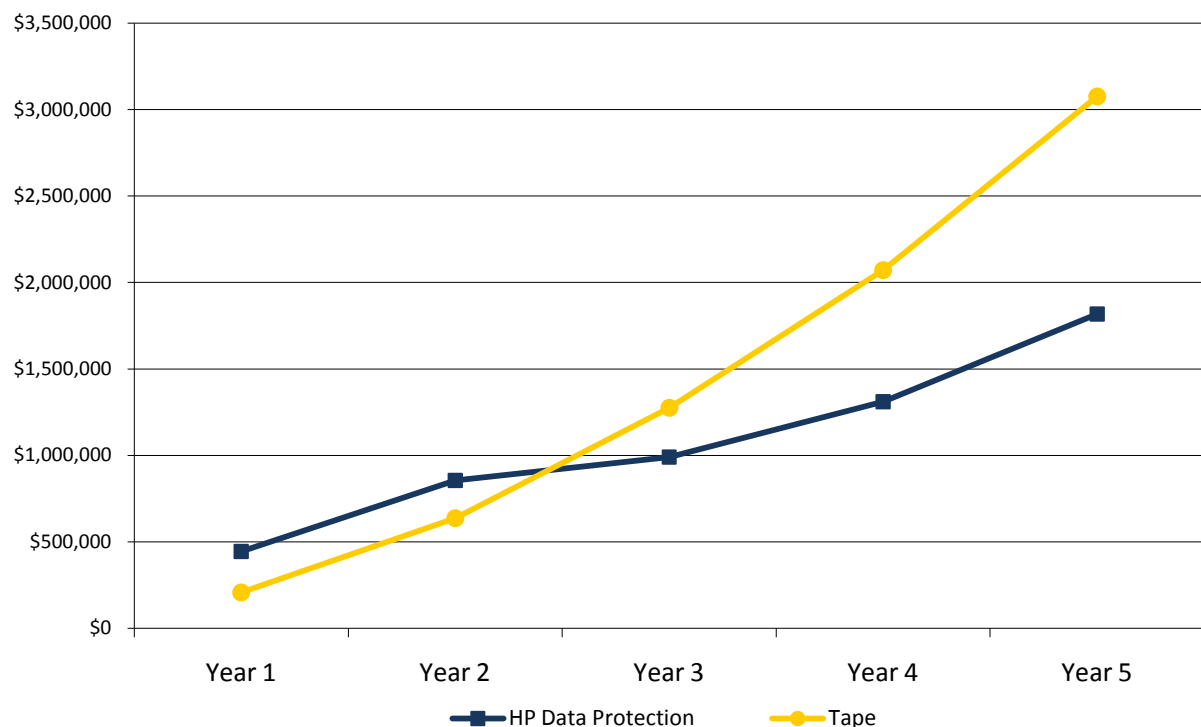


Figure 17 shows the total cost a hypothetical end-user would incur over five years when comparing a traditional tape environment to a backup to disk environment with replication and deduplication. The inflection point, where the disk environment becomes less costly than the tape environment, occurs just before the end of year two.

⁶ Assumptions and parameters can be found in the Appendix.

Table 2: Five Year Cost Breakdown by Category

	Capital Outlay	Administration	Maintenance	Power/Cooling/ Footprint	Total
Traditional Tape Environment	\$1,753,395	\$983,648	\$301,000	\$37,184	\$3,075,227
Disk with deduplication and replication	\$1,117,626	\$183,354	\$394,911	\$121,720	\$1,817,611

What the Numbers Mean

- The total cost of ownership of tape alone is roughly 69% higher than HP disk-based data protection with deduplication and replication.
- Data Protector software provides significant savings due to licensing based on data stored as opposed to data protected. Deduplication reduces licensing costs.
- The tape solution is more expensive in part due to the cost of acquiring tape media and the added complexity of managing the distributed tape infrastructure.
- While eliminating daily off-siting of tapes represents significant savings, tape is still the most cost-efficient method for long term archive of backups, and most organizations will replicate deduplicated data to a remote site for copy to tape for this purpose.

Why This Matters

Until recently, extending the benefits of a D2D2T protection strategy to remote and branch offices has been impractical. The cost of disk and WAN bandwidth for remote offices often can't be justified. If a disk-based storage system is used for replication, the backup software can't keep track of where the copies reside. Remote offices typically do not have the experienced IT staff needed to effectively administer tape or disk.

With HP StorageWorks D2D reducing the cost of disk capacity and WAN bandwidth and HP Data Protector to manage the data protection environment, ESG Lab has confirmed that HP's Data Protection solutions can extend the benefits of a centrally managed D2D2T strategy to an entire organization.

ESG Lab Validation Highlights

- ☑ Within five minutes of sitting down at the keyboard, ESG Lab was running a full backup of the first server to the HP VLS9000.
- ☑ The VLS 9000 demonstrated near linear performance scalability, achieving an impressive 4 TB/hour with a two-node configuration.
- ☑ The HP D2D Backup System achieved 81% bandwidth efficiency, transferring a 5 GB incremental backup across a 2 Mbit/sec simulated WAN link in just under 1 hour and 18 minutes for an effective throughput of more than 11 Mbit/sec. Replicating a second full backup yielded an even more impressive 90% bandwidth efficiency, transferring a 54 GB incremental backup across the same 2 Mbit/sec simulated WAN link in just over 6 hours and 42 minutes, for an effective throughput of more than 22.9 Mbit/sec.
- ☑ The HP Data Protection solution suite demonstrated faster, more reliable backups and restores at a significantly lower total cost of ownership than a tape environment.

Issues to Consider

- ☑ As with all VTLs today, when a cartridge is deleted or expired in a backup application, space on the VLS is not reclaimed until the cartridge is deleted, expired, or overwritten via the VLS management application. Integration with backup applications to automatically trigger a delete or expire in the VLS when a cartridge is expired in the backup application would be a useful enhancement.
- ☑ While the VLS, D2D, and Data Protector all have easy to use management interfaces, a single “manager of managers” that integrated all three products and provided an overall view of an entire enterprise environment would be of great value to administrators.
- ☑ While ESG is confident that one or more HP StorageWorks D2D Backup Systems can be used to meet the performance needs of a mid-sized organization, D2D systems with more capacity and horsepower could reduce cost and complexity within larger mid-sized organizations. HP has advised ESG that the new line of D2D Backup Systems, announced in June 2010, has been designed with these considerations in mind.

The Bigger Truth

ESG Lab conducted its first hands-on testing of the Hewlett Packard's enterprise VTL, the VLS 6000, in 2006 and then validated the VLS 9000 in 2008. Testing and discussions with end-users have confirmed that HP's disk-based backup solutions fit seamlessly into existing backup environments while providing dramatic performance and capacity reduction benefits compared to legacy tape-based methods. The HP StorageWorks D2D, aimed at delivering deduplication and WAN efficient replication to smaller, remote offices, completes a comprehensive, enterprise wide edge-to-core data protection architecture, managed by HP Data Protector, that goes beyond disk-based backup.

During this independent lab validation, ESG Lab confirmed HP's edge to core capabilities in support of large enterprises as well as deduplication support across disk, replication, and tape throughout the product line. HP's comprehensive offering with capacity efficient pricing provides high performance data deduplication technology to deliver dramatic disk capacity savings while offering scalable, predictable performance.

A modest two-node VLS configuration tested by ESG Lab was able to back up at a sustained 4 TB/hour. Based on the nearly linear scalability observed by ESG, an eight-node system should be able to protect data at 16 TB/hour. Easy to navigate, web-based management enabled ESG Lab to manage backups for a remote office from creation, through replication, and finally to tape—all using HP Data Protector software—giving rise to an edge-to-core data protection strategy covering remote and branch offices as well as multiple data centers. Direct attach to tape capability enables enterprises to meet offsite and deep archive requirements using familiar tools and techniques while keeping tape copy traffic off the SAN.

HP Data Protector software exemplifies the company's depth and breadth of end-to-end solutions for backup and recovery encompassing disk and tape. It should bring significant value to customers grappling with the challenges associated with cost-effective management of their data protection resources.

In essence, deduplication has become a crucial component of disk to disk backups, but when considering competing methods for implementation, customers should consider the tradeoffs and what's best for their organization: ease of implementation, cost, and bandwidth all play an important role.

ESG Lab believes that the combination of enterprise class performance and scalability—along with comprehensive storage management software and services—provides a unique approach for optimizing data protection and recovery strategy in the enterprise. Hewlett Packard customers can now retain more data for fast and reliable restores and longer retention periods while minimizing impact on backups with accelerated deduplication. Combined with the consolidated data management provided by Data Protector, customers have a wide choice of configurations which can be used to dramatically increase the role of disk in the protection of critical data.

Appendix

Table 3. ESG Lab Test Bed

Hardware	
HP StorageWorks VLS 9000	Firmware: 3.3-02
HP StorageWorks D2D 4009fc	Firmware: 101.041
Software	
HP Data Protector	Version: 6.1.1
WAN Simulator	
Network Nightmare	gigEnn-OC12
Total Cost of Ownership Assumptions	
Power	\$.0937/kW-hr
Data center floor space	\$40.00/Sq. Ft.
Hourly loaded labor rate for a backup admin	\$45/hour
Total working hours per year per backup admin	2,000 hours
Annual growth rate in required labor cost	10%
Data Center	
Capacity	25 TB
Backups	Weekly Full, Daily Incremental
Retention (tape)	30 Days onsite, 1 year offsite
Retention (disk)	90 Days onsite, 1 year offsite
Backup window	8 hours for a full Backup
Compressibility of data:	2:1 compression, 10:1 Deduplication over time
Incremental backup size	10% of production capacity
Data set growth rate	40% Annually
Full Restores required annually	Two (25 TB)
Partial restores required weekly	Five (>2.5 TB)
Tape infrastructure	Ultrium 920 LTO3 tape drives in MSL 2024 tape library
Remote Offices	
Capacity	50 GB
Backups	Weekly full, daily incremental
Retention (tape)	7 Days onsite, 1 year offsite
Retention (disk)	90 Days onsite, 1 year offsite
Backup window	2 hours for a full backup
Compressibility of data:	2:1 compression, 10:1 deduplication over time
Incremental backup size	10% of production capacity
Data set growth rate	40% annually
Full restores required annually	Two (50 GB)
Partial restores required weekly	Five (>5 GB)



Enterprise Strategy Group | **Getting to the bigger truth.**