

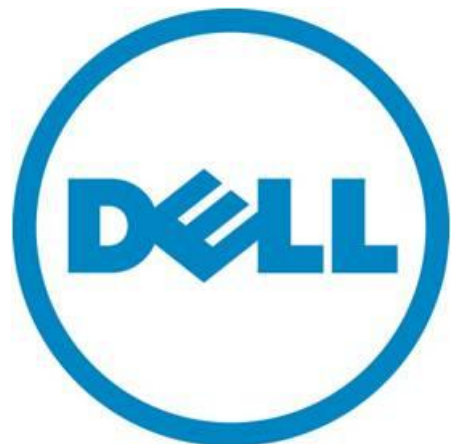
Dell SharePoint Infrastructure Optimization with AvePoint

A Dell | AvePoint Technical White Paper

Dave Jaffe
Solution Architect
Dell Solution Centers

Dave Zenz
Dell SharePoint Solutions Marketing

Dan Holme
Chief SharePoint Evangelist
AvePoint Inc.



Executive Summary

Dell's SharePoint Infrastructure Optimization solution takes a comprehensive Microsoft® SharePoint® infrastructure approach which consolidates SharePoint unstructured content, backup, and archive datasets on the Dell™ DX Object Storage Platform, thus improving end user performance, enabling better manageability, reducing capital and operational expenditures and increasing SharePoint service quality.

In this white paper a solution overview is provided, followed by details of the solution stack which includes components of AvePoint®'s DocAve® Software Platform and Dell hardware components. Finally, the set of functional and performance tests that were used to verify the solution are described, along with results.

After externalization of large content such as Word documents, PDFs, etc., from the SharePoint SQL Server database to the Dell DX Object Storage Platform using the DocAve Extender component, backing up a SharePoint site collection was 29x faster than backing up the non-externalized site collection, and restoring it was 32x faster. Routine SharePoint operations (browse, search, upload, and download) were 17% - 40% faster once this large content was moved to the external storage.

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

© 2011 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell. Dell, the DELL logo, and the DELL badge are trademarks of Dell Inc. Intel and Xeon are registered trademarks of Intel Corp. Windows and SharePoint are registered trademarks of Microsoft, Inc. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Dell Inc. disclaims any proprietary interest in trademarks and trade names other than its own.

November 2011

Contents

Executive Summary	2
Contents.....	3
The Dell SharePoint Infrastructure Optimization Solution	4
Dell DX Object Storage Platform.....	6
Solution Software Data-Control Flow	7
AvePoint’s DocAve Software Platform	8
Solution Hardware Configuration	10
Solution SharePoint Configuration	12
SharePoint Site Collections	12
DocAve Functionality Testing.....	14
Extender.....	14
Archiver.....	14
Backup/Restore	14
Extender Impact on SQL Backup and Restore Performance	16
SharePoint Load Performance with Extender.....	17
Acknowledgements	18

Tables

Table 1. AvePoint’s DocAve Software Platform Components - DocAve Manager.....	8
Table 2. AvePoint’s DocAve Software Platform Components - DocAve Agent.....	9
Table 3. Files in each SharePoint site used in tests.....	13
Table 4. SharePoint Site Collections Created	13
Table 5. Extender Test - 200 GB Site Collection.....	14
Table 6. Archiver Test - 10 GB Site Collection.....	14
Table 7. Backup and Restore of 10GB Site Collection	15
Table 8. Backup Performance with and without Extender.....	16
Table 9. Restore Performance with and without Extender	16
Table 10. Microsoft SharePoint Capacity Planner Heavy Collaboration Workload	17
Table 11. Extender Load Test Results.....	17

Figures

Figure 1. Detailed Data-Control Flow	7
Figure 2. Dell SharePoint Optimization with AvePoint’s DocAve Platform Reference Architecture ..	10
Figure 3. SharePoint Roles	12

The Dell SharePoint Infrastructure Optimization Solution

SharePoint is a multi-faceted platform that is being utilized at an increasing rate by corporations as a mechanism to manage and gain value from enterprise content. The growth of this content in the form of unstructured data and the increased reliance on SharePoint is challenging IT to adequately architect the underlying infrastructure to match and grow with business usage. Business as usual is not going to address SharePoint requirements especially considering that the typical backup/restoration requirements commonly associated with SQL Server usage are inefficient.

Organizations are realizing storage and data management must be brought into their main SharePoint operations environment to provide for unified SharePoint operations to reduce capital costs, operating expenses, and outage risk. What is required is a comprehensive SharePoint infrastructure approach which can seamlessly grow with the business, support tier 1 applications and their SLA's, improve end user performance, improve manageability, and consolidate SharePoint unstructured content, backup, and archive datasets.

Dell's SharePoint Infrastructure Optimization solution has the capability to interoperate and optimize existing SharePoint deployments seamlessly with no disruption to the applications or end users. As well leveraging the differentiated DX Object Storage Platform to consolidate SharePoint back-end data allows IT to scale the infrastructure according to the business demands.

The Dell DX Object Storage Platform and AvePoint's DocAve Software Platform can deliver a comprehensive, fully integrated solution for complete SharePoint Enterprise Content Lifecycle Management and Migration that can include all or some of the components based on the enterprise's needs.

- Simplify the migration process with the DocAve Migration Solution and the DX Object Storage Platform
 - A highly efficient and cost-effective solution to migrate existing enterprise content, Exchange Public Folders, File Systems and previous versions of SharePoint into SharePoint 2010.
- Improve SQL Performance and reduce SQL capacity requirements by as much as 90+% effectively eliminating backup and restoration challenges with DocAve Extender and the DX Object Storage Platform
 - Offloads Binary Large Objects (BLOBs) data from SQL Server databases to the DX Object Storage Platform - leveraging Microsoft's fully supported External BLOB Storage (EBS) or Remote BLOB Storage (RBS) APIs - to improve SQL performance while providing for seamless end-user access and interaction. To better understand the potential benefits and considerations for incorporating BLOB externalization into SharePoint content storage strategies, please reference [“Optimize SharePoint Storage with BLOB Externalization”](#) by SharePoint MVPs Dan Holme, Randy Williams, and Jeremy Thake.
- Automate SharePoint archive via business-driven rules to automatically apply content lifecycle management with DocAve Archiver and the DX Object Storage Platform
 - Empowers administrators to automate the process of offloading fixed or stale SharePoint content that would ordinarily take up valuable SharePoint resources to the easily manageable peer-scaling DX Object Storage.

- Robust data protection and quick restore with DocAve Backup and Restore and the DX Object Storage Platform
 - Quickly restore lost or corrupt assets, including content externalized with the DX Object Storage Platform and DocAve storage management solutions, with all metadata and dependencies intact to help ensure minimal business disruption.

Dell DX Object Storage Platform

The DX Object Storage Platform is a new generation of storage platform that combines ease of management with scalability and resilience. It's unique manner of storing data, be it SharePoint content, a backup dataset, or an archived SharePoint site as an object and not a collection of bytes, provides surprising ROI and TCO benefits that can be breakthroughs for IT departments.

The capabilities of the DX Object Storage Platform include:

Ability to Scale with Ease

- One storage pool with a global address space to consolidate SharePoint BLOBs, backup, and archive datasets
- Scale SharePoint to multi-Petabyte's - seamlessly with no service disruption
- All SharePoint capacity is virtualized and all data is automatically thin provisioned

Powerful Data and Storage Management Services

- Self-Managing - SharePoint access is inherently and automatically load balanced
- Self-Healing - all SharePoint content is automatically verified for consistency and repaired if necessary
- Policy based transactional replication for SharePoint BLOBs, no RPO scheduling required

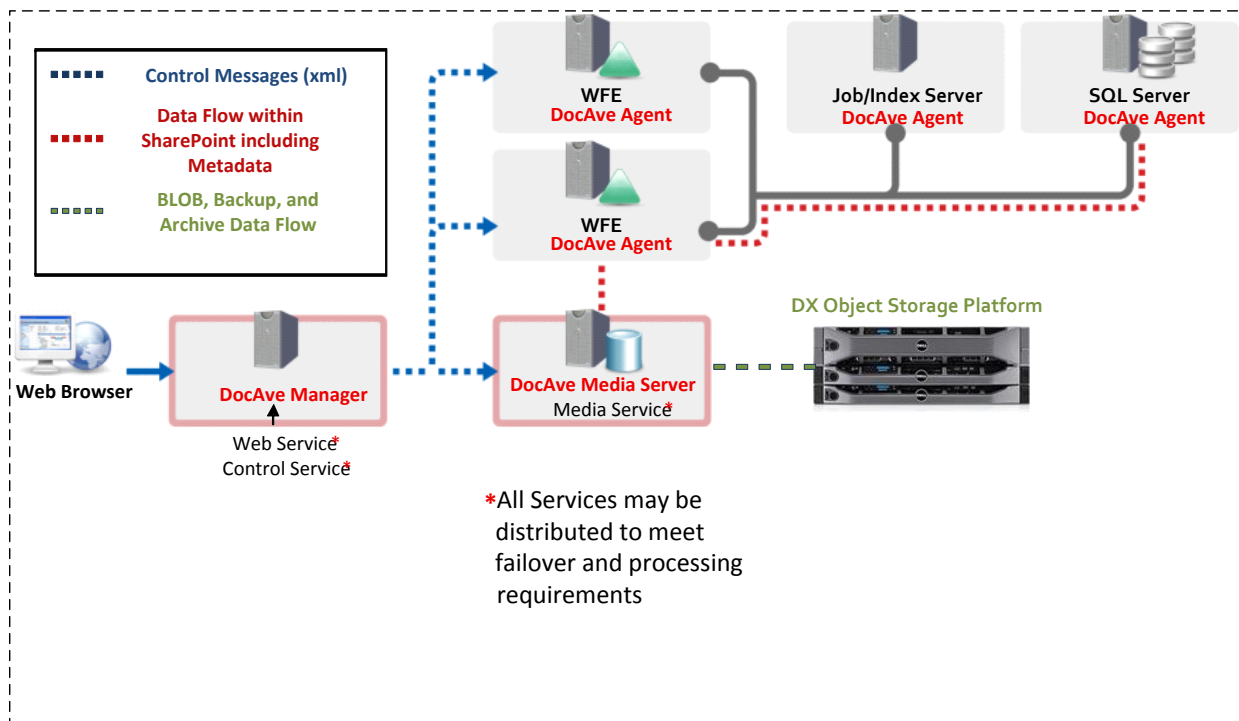
Future Proofed Design

- No database complexity, no file system, or LUNs to manage
- Optimized data protection for SharePoint content without the complexities of RAID
- No SharePoint application connection limit or application connection state to manage
- Storage technology and protocol impacts are removed from the SharePoint application and IT environments

Solution Software Data-Control Flow

The resultant Data Flow after the deployment of the SharePoint Infrastructure Optimization solution is enhanced by offloading the BLOB, backup, and archive flows from the main SharePoint environment. This augmented Data Flow serves to provide a parallelism of access to SharePoint, both for the end user as well as IT services, and also allows the optimal operation of the underlying SQL Server.

Figure 1. Detailed Data-Control Flow



AvePoint’s DocAve Software Platform

AvePoint's DocAve Software Platform is a truly integrated platform for SharePoint management, protection, and storage optimization, allowing organizations to scale SharePoint's architecture while simplifying SharePoint management. DocAve's uniquely distributed architecture not only allows for unmatched scalability, but also provides unified storage management of global SharePoint content on the DX Object Storage Platform.

DocAve provides a fully integrated and scalable software solution for organizations to streamline and simplify tasks associated with managing SharePoint environments, help comply with stringent SLAs including aggressive RTOs and RPOs for protecting SharePoint content, improve SharePoint performance for global users, optimize SQL resources, and can reduce risks associated with managing complex SharePoint environments by:

- Providing the tools necessary to streamline and automate common, yet time-consuming tasks that burden IT administrators, such as daily permissions management.
- Giving IT Managers insight into how SharePoint is being leveraged by end-users, so IT Managers can make better decisions on how to structure, manage, and grow SharePoint environments.
- Optimizing SQL resources to improve performance for end-users, and can lower total cost of ownership
- Providing IT Managers with robust solutions to quickly recover accidentally deleted or corrupted SharePoint content in order to help IT comply with stringent corporate SLAs and minimize business disruption.
- Allowing architects to expand SharePoint's storage and server architectures, improving quality of service for all global users while keeping costs in check.

The DocAve Software Platform components are shown in Table 1 and Table 2.

Table 1. AvePoint’s DocAve Software Platform Components - DocAve Manager

Component	Role
DocAve Manager	Web Service - Represents the access point to the DocAve GUI, allowing users to interact with the software. The web service can be installed on multiple machines and allows administrators to access the DocAve GUI using a web browser.
	Auditing Service - Monitors SharePoint activities and returns data to the Control Services for processing.
	Control Service - Manages all DocAve operations. All agent machines use the control service port and address to communicate with the manager, This service can be run on a server cluster to ensure high availability.
	Media Service - Manages all data storage for the DocAve platform. These can be installed on multiple machines and can manage multiple Logical Drives. Using multiple media agents will allow for high availability and load balanced access to the data storage locations.
	Search Service - Manages all data by generating a full text index.
	Report Service - Manages all SharePoint data collection and management for the DocAve Report Center. This must be configured with the associated SQL databases and SQL Report Services.

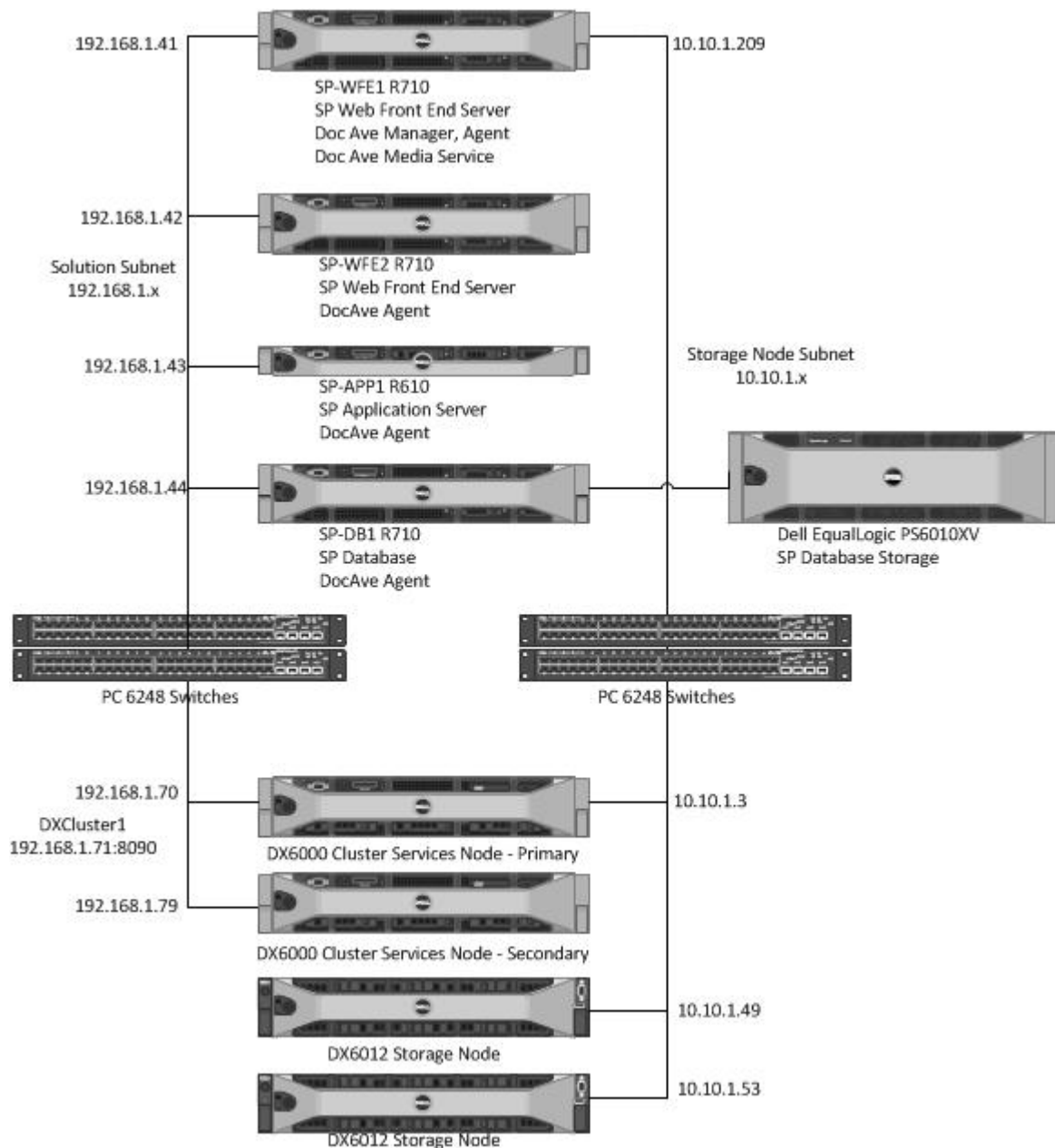
Table 2. AvePoint’s DocAve Software Platform Components - DocAve Agent

Component	Role
DocAve Agent	Storage Optimization - Archiver, Content & Media Library Connectors, Extender
	Data Protection - Granular Backup, High Availability Sync, Platform Recovery
	Administration - Administrator, Content Manager, Deployment Manager, Replicator
	Reporting - Report Center for Infrastructure Monitoring, Usage Analysis, and Alerting
	Compliance - Auditor, eDiscovery, Vault
	Migration - Exchange Public Folder, File Share, EMC Documentum, eRoom, Livelink, Lotus Notes & QuickPlace, SharePoint, Stellent, Vignette Migrators

Solution Hardware Configuration

The Dell SharePoint Infrastructure Optimization with AvePoint's DocAve Platform Reference Architecture is based on the Medium SharePoint 2010 Farm, supporting up to 5,000 users, as defined by Dell's Global Solutions Engineering team (for details see [Microsoft SharePoint Server 2010: Designing and Implementing a Medium Farm](#)). SharePoint is implemented in four servers: two SharePoint Front End Web Servers for high availability, a SharePoint Application Server, and a SharePoint SQL Server Database server. (For full high availability, a second Application Server can be added and a second SQL Server Database server can be added to form a SQL Server Failover Cluster).

Figure 2. Dell SharePoint Optimization with AvePoint's DocAve Platform Reference Architecture



Dell SharePoint Infrastructure Optimization with AvePoint

To the SharePoint infrastructure is added a highly available DX Object Storage Cluster, consisting of primary and secondary DX6000 Cluster Services Nodes, and two DX6012 Storage Nodes with 12 1 TB disks each.

The SharePoint servers and DX6000 Cluster Services Nodes are connected via a public network subnet (192.168.1.x). Additionally, a private network for the DX Object Storage Nodes (10.10.1.x) connects the DX Object Storage Cluster to the server running the DocAve Media Services (SP-WFE1 in this case).

Solution SharePoint Configuration

SharePoint roles are assigned to the various servers to maximize performance and availability. Looking specifically at the Search application (see Figure 3), the Admin and Crawl components are assigned to the application server while the search Index is spread across both web front end servers, with each web front end server's Index portion mirrored on the other web front end server. This provides higher performance (by spreading the work across multiple servers) as well as high availability. Additional web front end servers can be added for more Index servers and additional application servers can be added to provide additional search Crawlers, as well as to provide high availability. The two web front end servers are load balanced using Microsoft Network Load Balancing.

Figure 3. SharePoint Roles

Search Application Topology

Category	Server Name	Status
<div style="display: flex; justify-content: space-between; align-items: center;"> Current ▾ Modify </div>		
Admin		
Administration Component	SP-APP1	Online
Crawl - SP-DB1\Search_Service_Application_CrawlStoreDB_baa32f9d322f4d00b517c65355b5dbfe		
Crawl Component 0	SP-APP1	Online
Databases		
Administration Database : Search_Service_Application_DB_32d39fc7d4e247799c64474df9621f7d	SP-DB1	
Crawl Database : Search_Service_Application_CrawlStoreDB_baa32f9d322f4d00b517c65355b5dbfe	SP-DB1	
Property Database : Search_Service_Application_PropertyStoreDB_42dbe644b5644f2997ca68a7e70520cd	SP-DB1	
Index Partition - 0 - SP-DB1\Search_Service_Application_PropertyStoreDB_42dbe644b5644f2997ca68a7e70520cd		
Query Component 5	SP-WFE2	Online
Query Component 4	SP-WFE1	Online
Query Component 3	SP-WFE2	Online
Query Component 6	SP-WFE1	Online

SharePoint Site Collections

To facilitate testing, an automated method was developed to quickly populate large SharePoint site collections. First a content database with one site collection was created using the SharePoint administrator web site. Then a PowerShell script (available from authors) was run to create and load a specified number of team sites within that site collection. The content for each site is the same and consists of approximately 1 GB of typical document files: Microsoft Word documents, PDFs, text files, pictures (see Table 3 for details). For uniqueness each file has the site index number appended to the filename, for example QuickStart_Installation_001.pdf.

Table 3. Files in each SharePoint site used in tests

Number of files	173
Size range of files	21 B - 554 MB
Number of files larger than 1 MB	24
File types	.docx, .pdf, .txt, .bmp, .zip

The site collections built for testing are shown in the next table. Two large (200GB) site collections were built for testing the Remote BLOB Storage (RBS) feature of SQL Server as implemented by DocAve Extender. Two smaller (10GB) sites were built for testing DocAve Archive and Backup/Restore. The content databases holding the large site collections are stored on their own 400GB volumes on the Dell EqualLogic storage. The smaller site collections' content databases are stored on a third volume.

Table 4. SharePoint Site Collections Created

Site Collection Name	SQL Server Content Database	No of Sites	Files per Site	Size of each Site	Total Site Collection Size (approx.)	Test
SC1	WSS_Content_1	200	173	1 GB	200 GB	BLOBs remain in SQL, Crawled
SC2	WSS_Content_2	200	173	1 GB	200 GB	BLOBs Extended, DB Shrunk, Crawled
SC3	WSS_Content_3	10	173	1 GB	10 GB	Archived, DB Shrunk
SC4	WSS_Content_4	10	173	1 GB	10 GB	Backed up/Restored

DocAve Functionality Testing

Extender

The Remote BLOB Store (RBS) feature of SQL Server 2008 can improve SharePoint functionality immensely by removing the binary large objects (BLOBs) such as large PDF files from the SQL Server database and storing them on external storage, replacing them with stubs. DocAve software contains an RBS provider through their Extender functionality, which has been certified with the Dell DX Object Storage Platform.

To test the Extender functionality and simulate BLOB removal from an existing SharePoint site, one of the large (200GB) site collections, SC2, had all BLOBs larger than 1MB externalized to the DX Object Storage Cluster and then its content database was compacted using the SQL Serve Shrink function. The results are shown in Table 5. In just over 7 hours, 4,800 BLOBs (24 per site) were moved from SQL Server to the DX Object Store Cluster storage. Upon completion, the content database was compacted to 6.5 GB from its original 206 GB, a reduction of 97%.

Table 5. Extender Test - 200 GB Site Collection

Site Collection Size	Time for Extender Operation	Number of BLOBs Extended	Database Size after Shrink
206 GB	7 hours 22 min	4800	6.5 GB

To demonstrate that externalizing the BLOBs did not impact SharePoint functionality, both the Extended site collection (SC2) as well as the non-Extended site SC1 were crawled for search indexing. The crawls on both site collections took the same amount of time (1 hour, 3 minutes) and generated the same search results.

In the next section the performance impact of reducing the content database size on SQL Server backup and restore with Extender will be shown. Load tests in the final section demonstrate the impact of Extender on typical end user SharePoint operations.

Archiver

With DocAve Archiver objects meeting certain criteria - when last accessed, size, etc. - can be moved from SharePoint to the DX Object Storage Platform and maintained there according to a specified retention policy. To test this, all objects greater than 1MB from the 10 GB site collection SC3 was archived. As seen in Table 6, the operation took 24 minutes, 240 files were archived, and the Content database ended taking up less than 1 GB after shrink.

Table 6. Archiver Test - 10 GB Site Collection

Site Collection Size	Time for Archive Operation	Number of Files Archived	Database Size after Shrink
10 GB	24 min	240	.35 GB

Backup/Restore

DocAve Granular Backup and Restore provides for full or incremental backup of any specified sites, site collections, or content databases, to the DX Object Storage Platform. To test this a full backup of a 10 GB site collection, SC4, was created and then restored. Results are in Table 7.

Table 7. Backup and Restore of 10GB Site Collection

Site Collection Size	Number of Items Backed Up	Backup Time	Restore Time
10 GB	2642	22 min 55 sec	17 min 39 sec

Extender Impact on SQL Backup and Restore Performance

To quantify the impact of BLOB externalization on SQL Server database performance, SQL Server backups and restores were performed on the content databases of the two 200 GB site collections. As seen in the next two tables, backing up the Extended site collection was 29x faster than backing up the non-Extended site collection, and restoring it was 32x faster.

Table 8. Backup Performance with and without Extender

Site Collection Name	Status	Backup Time (sec)	Size of Backup File (GB)	Backup Performance Improvement over Not Extended DB
SC1	Not Extended	2811	207	-
SC2	Extended	98.3	6.42	29X

Table 9. Restore Performance with and without Extender

Site Collection Name	Status	Restore Time (sec)	Restore Performance Improvement over Not Extended DB
SC1	Not Extended	1533	-
SC2	Extended	47.2	32X

SharePoint Load Performance with Extender

To further evaluate the effect of Extender on SharePoint end user performance, a load test was built using Microsoft Visual Studio 2010 Test to simulate typical SharePoint use, and run against both the site collection with BLOBs still in SQL Server (SC1) and the site collection with BLOBs externalized to the DX Object Storage Platform store with DocAve Extender (SC2).

Four typical use cases - Browse, Download File, Upload File and Search - were modeled. Browsers involved users going to the various site home pages as well as certain list pages. Downloads and uploads involved the same set of files of varying sizes that were used to build each site. Search used a common set of technical search terms relevant to the file set. All tests were run against all 200 sites of each site collection.

The four use cases were run in a proportion to model the Microsoft SharePoint 2010 Capacity Planner (SCP) “Heavy Collaboration” usage profile, as shown the next table. In each test 1,000 simulated users were run for 60 minutes, with a 60 second think time inserted between operations, so that each user would issue roughly one request per minute, again according to the SCP.

Table 10. Microsoft SharePoint Capacity Planner Heavy Collaboration Workload

Operation	Mix
Browse	57
Download file	17
Upload file	9
Search	17
Total	100

The results of the test are shown in the next table. Using DocAve Extender to move BLOBs from SQL Server to the Dell DX Object Storage Platform results in performance improvements of from 17% to 40%.

Table 11. Extender Load Test Results

		SC1 - BLOBs remain in SQL Server			SC2 -BLOBs Extended to DX Object Storage			Extended Faster than Non-Extended
Request	Start Mix	Requests Completed	%	Average Test Time (s)	Requests Completed	%	Average Test Time (s)	
Browse	57	33,243	56.8%	0.91	33,705	57.1%	0.64	30%
Download	17	10,197	17.4%	1.26	10,035	17.0%	0.76	40%
Upload	9	5,146	8.8%	4.53	5,296	9.0%	3.42	25%
Search	17	9,913	16.9%	1.53	9,964	16.9%	1.27	17%

Acknowledgements

Dave Jaffe would like to thank Quocdat Nguyen and Ravi Chaganti of the Dell Global Solutions Engineering team for much SharePoint and PowerShell assistance, Omar Rawashdeh of the Dell Storage Engineering team for help setting up the DX Object Storage Platform with DocAve, and Roger Yu of AvePoint for much help with installing and running DocAve.