EMC²
where information lives®

White Paper

EMC Tiered Storage for Microsoft Office SharePoint Server 2007 BLOB Externalization Enabled by EMC CLARiiON, EMC Atmos, Microsoft Hyper-V, and Metalogix StoragePoint

Applied Technology

EMC Information Infrastructure Solutions

BROCADE

Abstract

This white paper demonstrates how EMC® Atmos® cloud-optimized storage, together with Metalogix StoragePoint®, provides a dynamic external BLOB store (EBS) solution for Microsoft® Office SharePoint Server (MOSS) 2007. Organizations will learn how to establish a highly scalable, effective approach for managing SharePoint data growth using the infinite storage capacity of cloud-based storage.

September 2010

Part number: H7048

Table of Contents

Executive summary

Overview

Many organizations depend on Microsoft® Office SharePoint Server (MOSS) 2007 for its dynamic content management and collaboration features. But at scale, SharePoint farms containing ever-expanding, multi-terabyte content databases pose a significant drain on existing storage.

Part of the difficulty lies in the current method most administrators use to manage unstructured content. Such content is stored as binary large objects (BLOBs) in the Microsoft SQL Server database itself rather than in a file system. This means that all SharePoint operations involving objects (including read, write, index, and backup) are bound by the performance and expense of the underlying SQL database system—which is not designed to manage unstructured files.

This white paper demonstrates how EMC® Atmos® and Metalogix StoragePoint® alleviate the constraints associated with managing SharePoint objects stored in a SQL Server database by externalizing the content to a distributed, scalable cloud storage platform while maintaining overall SharePoint performance and user experience.

Business case

On a daily basis, organizations are creating huge volumes of unstructured content within multiple organically grown, unmanaged, incompatible systems. This unstructured content varies, including documents, e-mail, video files, and web pages. The content is often disorganized, which can lead to inefficient information sharing and reduced employee productivity.

The infrastructure required to support such disparate content tends to cause physical server and storage sprawl, leading to additional data center demands. This results in increased energy consumption, a larger physical footprint, and high maintenance costs.

This white paper validates how organizations can:

- Grow the SharePoint farm to an enterprise-scale content management system versus limiting content databases to 100 GB (Microsoft's current recommendation for SharePoint 2007 environments reliant on SQL Server).
- Establish a reliable, highly effective method for managing future SharePoint data growth (as more content is added, more space is allocated from Atmos cloud storage).

Product overview

This white paper presents a validated method for eliminating the constraints commonly associated with maintaining traditional, SQL-based SharePoint content databases by using StoragePoint to move data to Atmos, a cloud storage platform.

Testing focuses on a virtualized Microsoft Office SharePoint Server 2007 enterprise farm environment, and a Microsoft Hyper-V infrastructure that consolidates multiple application, web, and database servers on a smaller set of physical servers. The success of this validation is driven by Atmos, which anchors the configuration, integrating a comprehensive set of advanced management features. The Atmos product line is designed to meet a range of IT infrastructure needs, including support for:

- A growing ecosystem of Atmos-powered service providers such as AT&T, and PEER 1
- Content rich web applications, archiving, and infrastructure as a service

Key results

This white paper demonstrates StoragePoint integration as a highly efficient method for moving SharePoint data to Atmos cloud storage. The design outlined in this white paper:

- Decreases the size of SharePoint SQL databases, and subsequently reduces maintenance windows.
- Frees BLOB content that would normally be constrained in a SQL database, using Atmos cloud storage to reduce the organization's overall storage levels and physical footprint without compromising performance or SharePoint workflow and processes.
- Demonstrates how to use StoragePoint for effectively managing the externalization of SharePoint BLOBs out of SQL Server content databases.

Introduction

Introduction to this white paper

This white paper includes the following sections:

Purpose

The purpose of this white paper is to document an enterprise SharePoint 2007 environment using external BLOB storage on a cloud infrastructure provided by Atmos. The document presents validated test results on the scalability and efficiency of the SharePoint farm before and after BLOB externalization.

Scope

This white paper presents design guidelines, best practices, and validated test results that were identified during the configuration and test phases of the Microsoft Office SharePoint Server 2007 BLOB externalization enabled by EMC CLARiiON, EMC Atmos, and Metalogix StoragePoint.

Audience

This white paper is intended for:

- EMC employees
- EMC partners
- Customers, including IT planners, storage architects, and administrators
- Field personnel, who are tasked with designing and implementing cloud storage in a virtualized SharePoint Server 2007 environment on EMC storage

Terminology This white paper includes the following terms.

Term	Definition
binary large object (BLOB)	A collection of binary data stored as a single entity in a database management system. Typically, BLOBs are multimedia objects (images and audio).
BLOB store	A framework for moving unstructured data from the SQL server to an external storage system, like EMC Atmos.
Cloud storage	Public or private enterprise-based storage infrastructure designed to help organizations automatically manage and optimize the distribution of rich, unstructured information across geographically dispersed locations.
Content database	Content can be separated into multiple content databases at the site collection level. A content database can include one or more site collections, but a single site cannot span multiple databases. Backing up and restoring sites take place at the content database level.
External BLOB store (EBS)	Microsoft released this application programming interface (API) to support external data stores. EBS intercepts the read/write activity directed at the SQL Content Server and dictates whether to store the data in the database or to redirect to external storage.
Non-externalized storage	Content resides on Tier 1 storage within the SQL database.
Subtenant	Subtenants are logical partitions of tenants that group together selected policies, data access, and reporting capabilities. Each subtenant has a unique set of users and processes a unique set of data.
Tenant	Logical units of data and resources that are associated with specific access nodes, security control, and storage policies. A tenant is configured with web-service or file-system access to specific nodes.
Thin LUNs	A logical unit of storage where physical space allocated on the storage system may be less than the user capacity seen by the host server.
Thin pool	A group of disk drives used specifically by thin LUNs. There may be zero or more thin pools on a system. Disks may be a member of no more than one thin pool.

Technology

Overview

The following sections identify and briefly describe the technology and components used in the configuration, including:

- Microsoft Office SharePoint Server 2007
- EMC CLARiiON CX4-480
- EMC Atmos WS2-120
- Metalogix StoragePoint
- Microsoft Hyper-V

Microsoft Office SharePoint Server 2007

Microsoft Office SharePoint Server 2007 is a server application that facilitates collaboration, provides content management features, and implements business processes. It provides an integrated platform to plan, deploy, and manage intranet, extranet, and Internet applications across and beyond the enterprise. SharePoint uses multiple servers in various roles to organize and deliver website collaboration and information sharing across organizations.

EMC CLARiiON CX4-480

EMC CLARiiON CX4-480 provides high-capacity networked storage that meets the needs of demanding online transaction processing (OLTP) workloads and large-scale e-mail environments. With the CX4-480, customers can scale seamlessly up to 471 TB of storage capacity and consolidate twice the workloads in one array as is possible with other storage providers.

EMC Atmos WS2-120

EMC Atmos is a cloud storage platform designed to help customers automatically optimize the distribution of rich, unstructured information across geographically dispersed locations and is managed using the Atmos GUI, command line, or API. Its globally distributed architecture and management capabilities make it ideal for protecting content at Internet scale. EMC Atmos provides the most extensive range of advanced management features for cloud storage, including:

- **Policy-based management:** Automatically distributes content based on business policy. The administrator-defined policies dictate how, when, and where the information resides.
- **Cloud-optimized protection:** Customers can control the protection, performance, and storage cost associated with the content using the Atmos GeoProtect feature that integrates:
  - Replication to create multiple copies of the object to one or more location.
  - Erasure coding to ensure that the content is always available by dividing objects into multiple segments and distributing them to one or more designated locations.

- **Multi-tenancy:** Enables multiple applications and groups of users to be logically segregated amongst the same shared infrastructure. Each application is securely partitioned and cannot access another application's data.
- **All-in-one data services**: Atmos data services include replication, compression, deduplication, and spin down.

EMC supports Atmos in three hardware configurations: WS2-120, WS2-240, and WS2-360. The WS2-120 and WS-240 provide a higher mix of system resources, while the WS2-360 provides the most storage capacity.

Microsoft Windows Server 2008 R2 with Hyper-V

Hyper-V is a hypervisor-based virtualization technology from Microsoft that makes it easier than ever to take advantage of the cost savings of virtualization through Windows Server 2008 R2.

Hyper-V enables customers to make the best use of server hardware investments by consolidating multiple server roles as separate VMs running on a single physical machine. It can also efficiently run multiple different operating systems – Windows, Linux, and others – in parallel, on a single server.

Metalogix StoragePoint

StoragePoint reduces the size of SharePoint content databases by relocating (externalizing) content BLOBs from the database and maintaining only the metadata within SQL Server. BLOBs are relocated to the storage area network (SAN), NAS, or cloud platforms where they can be partitioned by SLA, isolation requirements, or retention policies. The BLOB I/O is moved from SQL Server to the SharePoint web front ends (WFEs), which are easier to scale than SQL Server. Scalability is addressed by adding more web servers to the farm.

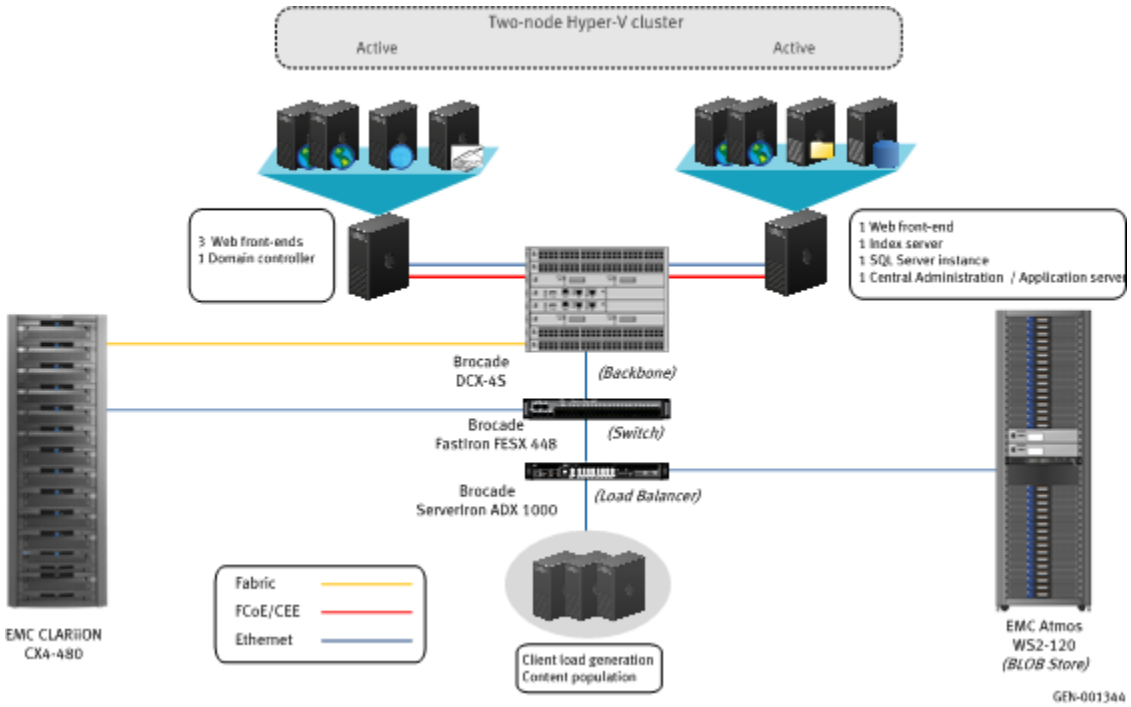
Environment profile

Overview

The following section identifies and briefly describes the technology components used in the environment.

Physical environment

The following diagram illustrates the overall physical architecture of the environment.



Hardware

The following table lists the hardware used in testing.

Equipment	Quantity	Configuration
Storage	1	EMC CLARiiON CX4-480 Code: FLARE® 29 Version: 04.29.000.5.003
Fibre Channel Switch	2	8 Gb/s Ethernet Class Fibre Channel switch (requires a minimum of 48 ports)
Brocade® PowerConnect B- DCX 4S network switch	1	Gigabit Ethernet network switch (requires a minimum of 32 ports)
Brocade FastIron Edge X448 Switch	1	Provides 10 Gigabit Ethernet uplinks
Brocade ServerIron ADX 1000 Load Balancer	1	Used for Atmos on-premise and SharePoint WFEs
Server, quad core, 2 socket with Xeon processor	2	16 logical CPUs, 2.26 GHz 72 GB RAM
Server, quad core, 2 socket with Xeon processor	1	16 logical CPUs, 2.26 GHz, 32 GB RAM
Brocade Dual HBAs	2	10 Gb/s Ethernet Adapter (PCI\VEN_1657, DEV_0014 and CC_020000)
Converged Network Adapter	2	10 GB FCoE

Software

The following table lists the software used in testing.

Item	Version	Configuration
Microsoft Office SharePoint Server 2007	SP2	Installed on web front end (WFE) servers and application servers.
Microsoft SQL Server 2008	64-bit Enterprise Edition SP1	Installed on database servers.
Windows Server 2008 Enterprise Edition R2	RTM	Installed on WFE servers, application servers, and database servers that are running as virtual machines (VMs).
Microsoft Hyper-V	2008 R2	Used as a hosting platform for all SharePoint-related services running on VMs.
Metalogix StoragePoint	2.1	Default settings applied.

Virtual allocation of resources

The following table lists the virtual allocation of hardware resources used in the environment.

VM role	Quantity	Configuration
WFE	4	4 CPUs, 2.26 GHz, 4 GB RAM
Application	1	2 CPUs, 2.26 GHz, 2 GB RAM
Index	1	4 CPUs, 2.26 GHz, 6 GB RAM
SQL	1	4 CPUs, 2.26 GHz, 24 GB RAM
DC	2	2 CPUs, 2.26 GHz, 2 GB RAM
Central Admin	1	2 CPUs, 2.26 GHz, 2 GB RAM

Disk types

The disk types used in the environment include:

- Dynamic virtual hard disks (VHDs)—Used for content data on thin LUNs.
- Fixed VHDs—Used for transaction logs, tempdb, and search indexes.

Storage design layout

Overview

The following sections detail the storage design guidelines that EMC followed during the build of this environment. The process is essentially the same as designing a physical environment, except that it requires a physical server for each VM.

Disk layout

The following image represents the physical disk distribution for the CX4-480 at the production site using 300 GB (268.4) Fibre Channel (FC) drives.

Slot	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Bus 0 Enclosure 0	RAID 5 -268.4 RAID Group - 160	RAID 5 -268.4 RAID Group - 160	RAID 5 -268.4 RAID Group - 160	RAID 5 -268.4 RAID Group - 160	RAID 5 -268.4 RAID Group - 160	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	RAID 1/0 -268.4 RAID Group - 162	RAID 1/0 -268.4 RAID Group - 162	RAID 1/0 -268.4 RAID Group - 161	RAID 1/0 -268.4 RAID Group - 161	RAID 5 -268.4 Thin Pool - 0
Bus 0 Enclosure 1	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4
Bus 1 Enclosure 0	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0
Bus 1 Enclosure 1	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0	RAID 5 -268.4 Thin Pool - 0
Bus 2 Enclosure 0	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	Free Disk -268.4	RAID 1/0 -268.4 RAID Group - 169	RAID 1/0 -268.4 RAID Group - 169	RAID 1/0 -268.4 RAID Group - 169	Free Disk -268.4	Free Disk -268.4	RAID 1/0 -268.4 RAID Group - 169
Bus 2 Enclosure 1	RAID 5 -268.4 RAID Group - 168	RAID 5 -268.4 RAID Group - 168	RAID 5 -268.4 RAID Group - 168	RAID 5 -268.4 RAID Group - 168	RAID 5 -268.4 RAID Group - 168	RAID 1/0 -268.4 RAID Group - 167	RAID 1/0 -268.4 RAID Group - 167	RAID 1/0 -268.4 RAID Group - 166	RAID 1/0 -268.4 RAID Group - 166	RAID 1/0 -268.4 RAID Group - 165	RAID 1/0 -268.4 RAID Group - 165	RAID 1/0 -268.4 RAID Group - 164	RAID 1/0 -268.4 RAID Group - 164	RAID 1/0 -268.4 RAID Group - 163	RAID 1/0 -268.4 RAID Group - 163
Bus 3 Enclosure 0	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18	Free Disk -917.18

GEN-001502

LUN distribution

The following table details LUN distribution in the test environment.

Note
Half of the content databases reside on SP A, and the remaining half resides on SP B. This design balances the load across the two storage processors (SPs).

SP	LUN number	LUN name	Capacity	Storage pool	RAID type
SP A	604	ContentDB Log SPA	268.371 GB	RG 161	RAID 1/0
	606	TempDB SPA	268.371 GB	RG 163	RAID 1/0
	608	SearchDB Data SPA	268.371 GB	RG 165	RAID 1/0
	610	SearchDB Log	268.371 GB	RG 167	RAID 1/0
	612	Query SPA	536.742 GB	RG 168	RAID 5
	613	Content Index	500 GB	RG 169	RAID 1/0
SP B	605	ContentDB Log SPB	268.371 GB	RG 162	RAID 1/0
	607	TempDB SPB	268.371 GB	RG 164	RAID 1/0
	609	SearchDB Data SPB	268.371 GB	RG 166	RAID 1/0
	611	Query SPB	536.742 GB	RG 168	RAID 5

Thin LUNs

The following table details thin LUN distribution in the test environment.

Folder name	Thin LUN number	Thin LUN name	User capacity (GB)	Consumes capacity (GB)	Storage pool	RAID type
SP A	602	ContentDB Data SPA	2048	1592.097	Thin pool 0	RAID 5
SP B	603	ContentDB Data SPB	2048	1591.097	Thin pool 0	RAID 5

Thin pool

The thin pool (identified as thin pool 0 in the test environment) is comprised of the following thin LUNs:

- ContentDB Data SP A
- ContentDB Data SP B

The following table details thin pool usage in the test environment.

RAID type	Total capacity (GB)	Consumes capacity (GB)	Percent full	Total subscribed capacity	Percent subscribed	Percent full threshold	Total thin LUNs
RAID 5	4280.261	3183.194	74.369%	4,096	95.695%	70%	2

SharePoint configuration in the test environment

Overview

The following sections detail the SharePoint environment profile.

Validated environment profile

The following table defines the SharePoint 2007 configuration.

Profile characteristic	Value
SharePoint farm scale	184,800
Portal type	Publishing/collaboration
SharePoint sites	30
Site collections	30
Estimated overall content size	3 TB
Number of content databases	30
Web applications	7
Data profile	100 GB random user data
Production drive size and speed	300 GB, 15k, RAID 5

SharePoint 2007 VM farm configurations

The SharePoint 2007 VM farm configurations consist of:

- **Four WFE VMs**—This resource distribution offers the best search performance and redundancy in a virtualized SharePoint farm. Because WFE and query roles are CPU-intensive roles, the WFE VMs are allocated using four vCPUs with 4 GB of memory.
- **Index server**—The index server is configured as the sole indexer for the portal along with a dedicated WFE role. This means that while the index VM is crawling for content, it can use itself as the WFE to crawl. This minimizes network traffic and ensures that SharePoint farm performance does not suffer.

  Four vCPUs and 6 GB of memory are allocated for the index server.
- **Application Excel servers**—Two vCPUs and 2 GB of memory were allocated for the Application and Excel servers as these roles require less resources.
- **SQL server**—Four vCPUs and 24 GB of memory are allocated for the SQL server VM as CPU utilization and memory requirements for SQL in a SharePoint farm are high. With more memory allocated to the SQL VM, the SQL server becomes more effective in caching SharePoint user data, leading to fewer required physical IOPS for storage and better performance.

SharePoint content databases

The following table lists the SharePoint content databases.

Content database	Database Size (GB)
Content_DB_01	109
Content_DB_02	110
Content_DB_03	109
Content_DB_04	109
Content_DB_05	109
Content_DB_06	110
Content_DB_07	110
Content_DB_08	109
Content_DB_09	109
Content_DB_10	113
Content_DB_11	110
Content_DB_12	109
Content_DB_13	111
Content_DB_14	111
Content_DB_15	111
Content_DB_16	109
Content_DB_17	111
Content_DB_18	111
Content_DB_19	111
Content_DB_20	110
Content_DB_21	111
Content_DB_22	110
Content_DB_23	112
Content_DB_24	115
Content_DB_25	111
Content_DB_26	111
Content_DB_27	115
Content_DB_28	115

Content database	Database Size (GB)
Content_DB_29	111
Content_DB_30	111

Networking

The following table shows how the network is divided into multiple segments to improve overall throughput by segregating competing traffic.

IP range	Network VLAN	Description
10.13.242.90	Default	Public network A
192.168.1.0/24	20	CSV
192.168.2.0/24	30	Cluster heartbeat network
192.168.3.0/24	40	Live migration
192.168.4.0/24	50	Simulated public network
192.168.13.0/24	612	Management

Atmos design considerations

Overview

Review this section for an overview of the configuration steps required to enable Atmos for StoragePoint, as follows:

- Set up resource management groups (RMGs)
- Create a tenant list
- Configure resource nodes
- Configure subtenants
- Set credentials

Prerequisites for Atmos

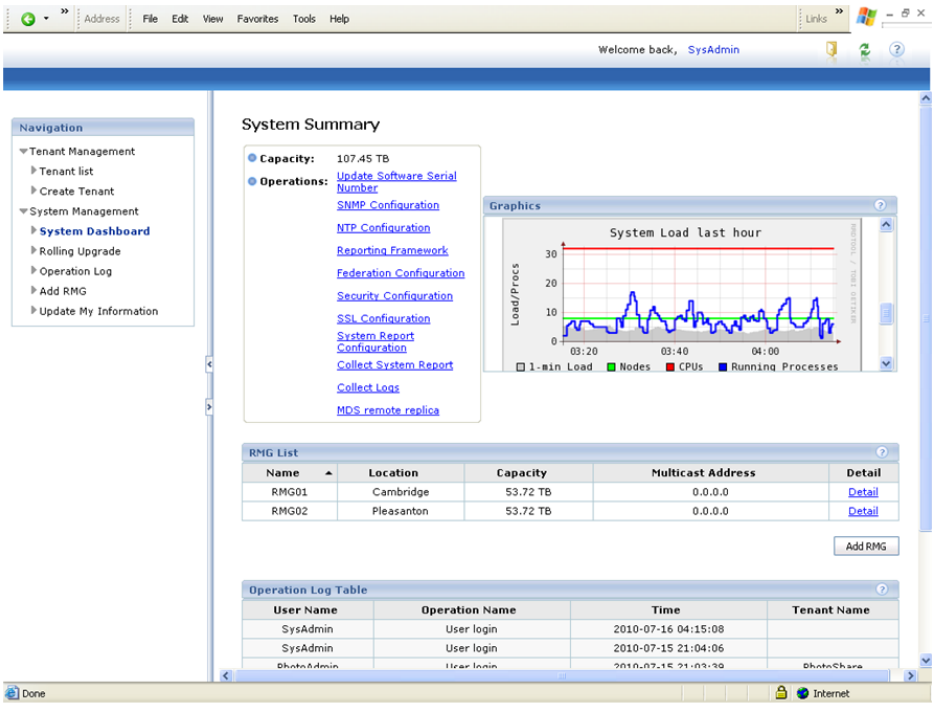
Follow these best practices before configuring Atmos in the environment:

- Implement a load balancer to distribute I/O between multiple Atmos nodes.
- Define multiple subtenants in order to quickly provide chargeback information based on line of business, department, or SharePoint site collection.

Note
For detailed information, including complete procedures, refer to the *EMC Atmos Administrator's Guide, Version 1.3.2*, Revision 1.3.0A. For detailed information, including complete procedures, refer to the *EMC Atmos Administrator's Guide, Version 1.3.2*, Revision 1.3.0A.

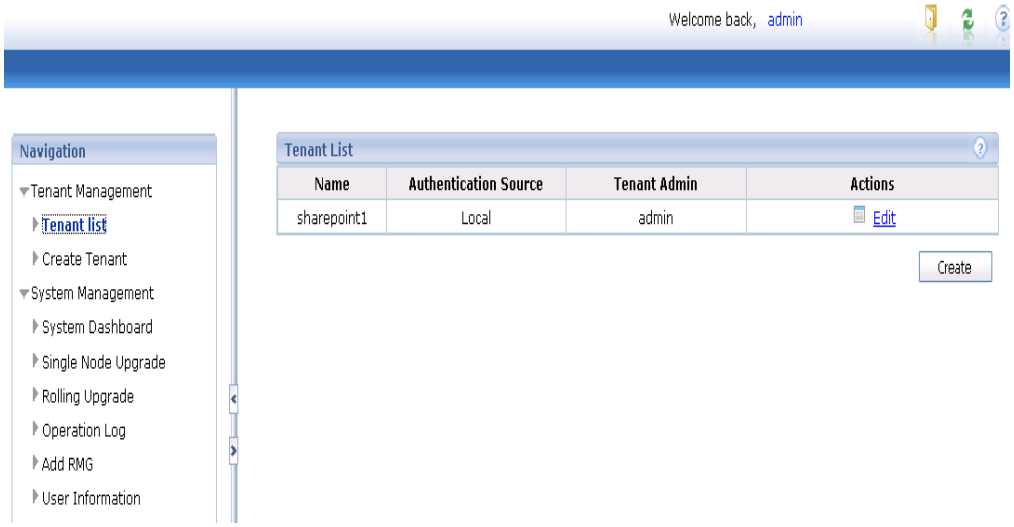
Set up resource management groups

A resource management group (RMG) is a collection of Atmos nodes, either physical or virtual, that share a single IP multicast domain. An RMG is typically, but not always, associated with a specific physical location. Atmos is configured with two RMGs (rmg1 and rmg2) as shown in the following image.



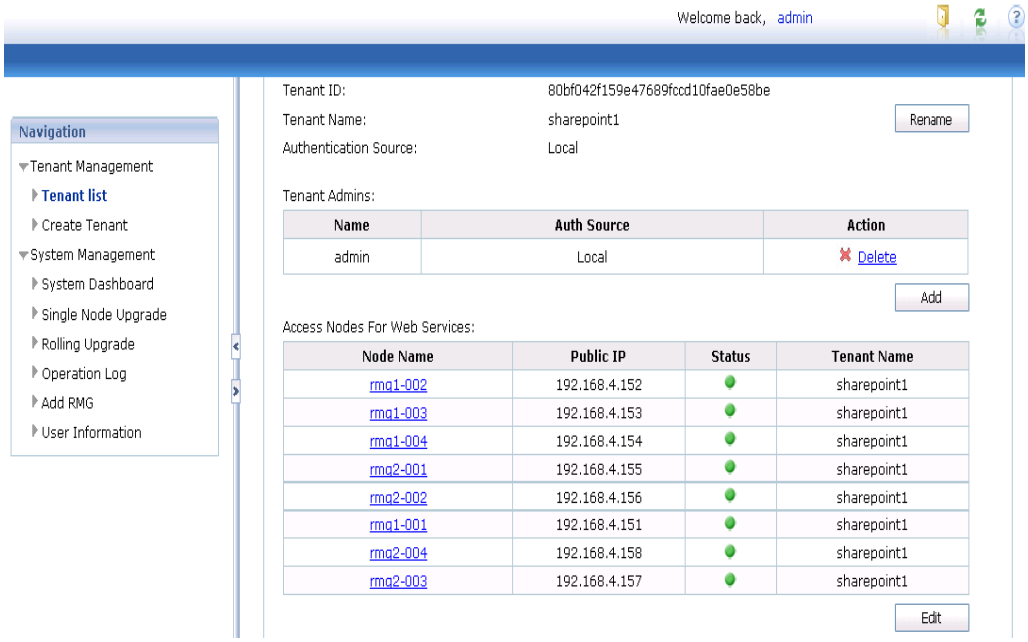
Create a tenant list

A tenant list is set up to generate the required credentials for the cloud BLOB externalization as shown in the following image.



Configure resource nodes

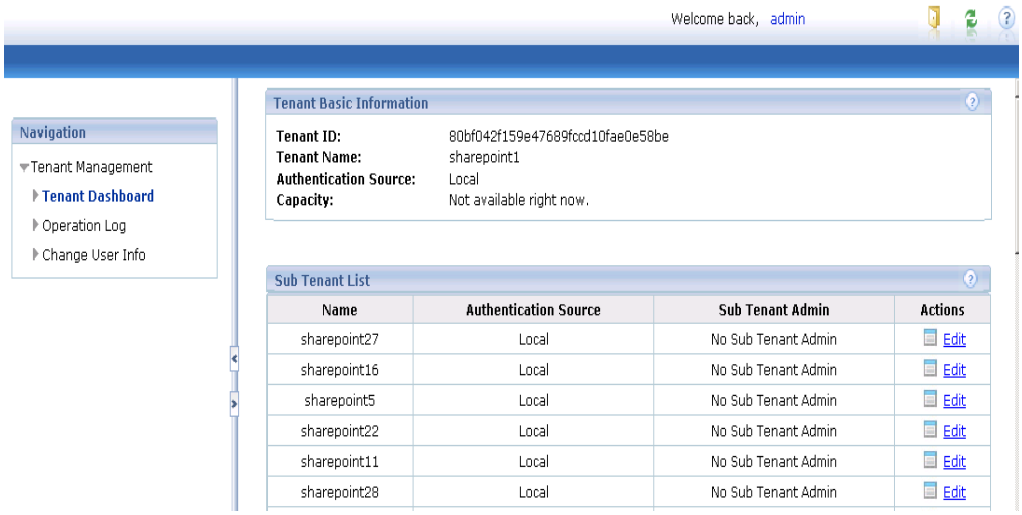
All resource nodes from Atmos are configured to allow the sharepoint1 tenant maximum performance on all Atmos nodes, as shown in the following image.



Configure subtenants

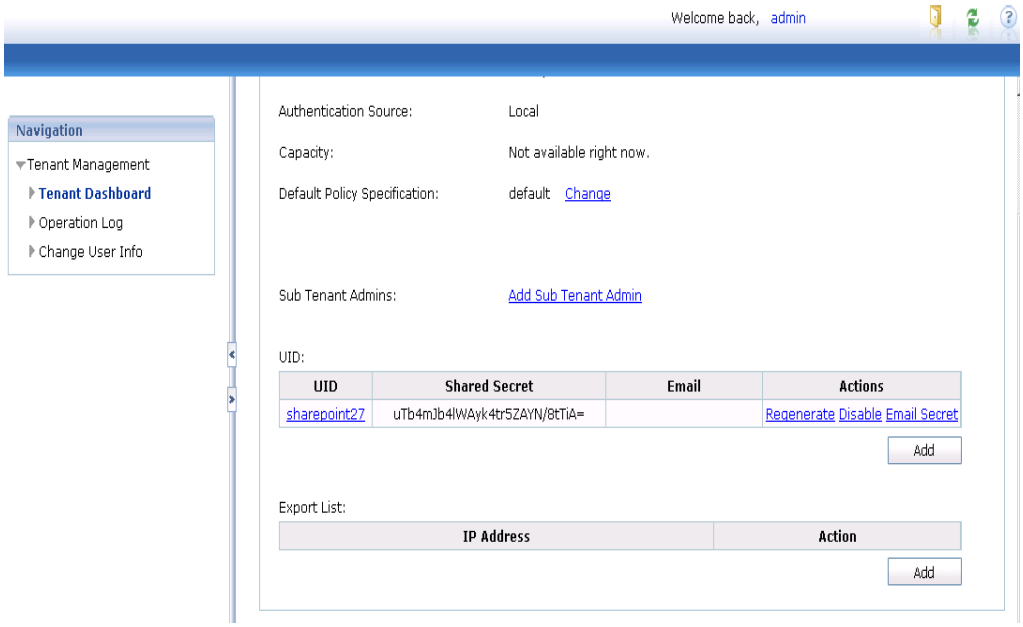
The test environment consisted of a 3 TB SharePoint instance running on the SQL Server. EMC created 30 content databases to support the SharePoint setup based on SQL Server best practices. This configuration required 30 subtenants for each of the content databases as shown in the following image.

For more information on SQL Server best practices, see http://technet.microsoft.com/en-us/library/cc850692(office.12).aspx.



Set credentials

Each subtenant has unique credentials that are needed to configure each content database with StoragePoint software. The credentials information (called Shared Secret) is taken from the details of each of the subtenants as shown in the following image.



StoragePoint design considerations

Overview

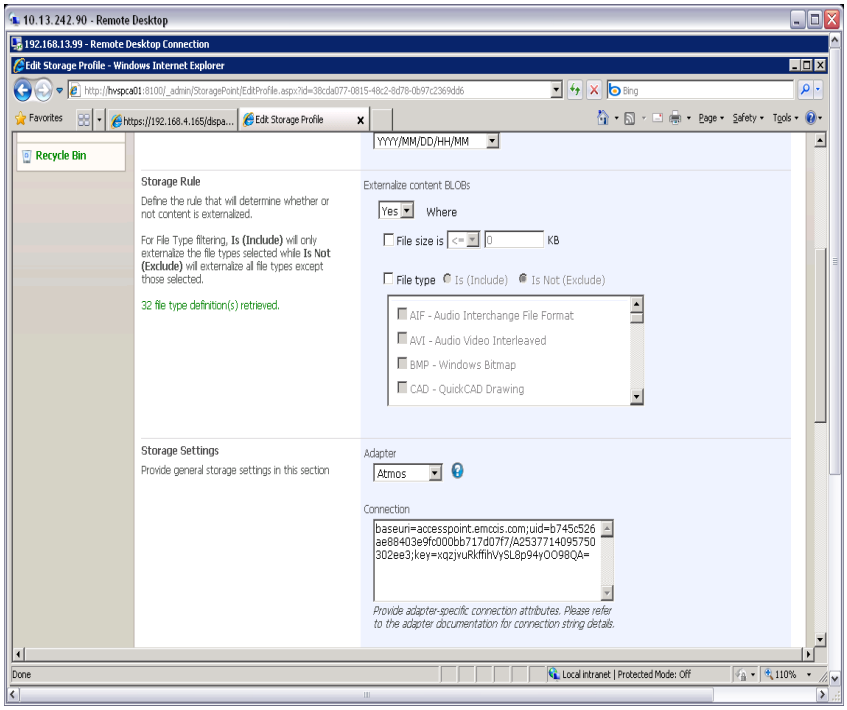
StoragePoint enables organizations to relocate (or externalize) content BLOBs to a location outside of the SQL database to a file system using SAN/NAS storage or to an object-based storage like EMC Atmos. As test results demonstrate, this allows database administrators to take full advantage of cloud-based storage and all it provides in terms of scale and high availability (HA).

Using StoragePoint to externalize content BLOBs

StoragePoint is integrated with SharePoint Server 2007 as a plug-in, and configured using the application's default settings. Each content database needs the credentials for the external cloud storage location. Configuring StoragePoint to work with Atmos is a simple, two-step process as described next.

For more detailed procedures, see the *Metalogix StoragePoint Installation and Administration Guide* at http://www.storagepoint.com/adminguide.pdf.

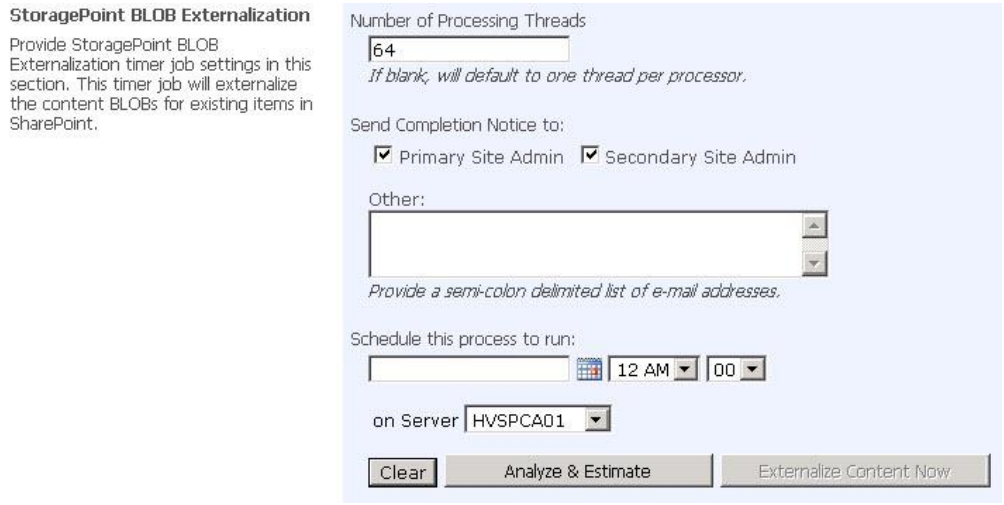
1. Navigate to the **Edit Storage Profile** tab in SharePoint Central Administration and set the **Storage Settings** parameters for Atmos, as shown in the following image.



Note
The connection string determines the target BLOB storage, the tenant, and shared secret key.

2. Next, set the number of processing threads for each content database, as shown in the following image.

Note
The parameter is set to 64 threads in the test environment to achieve minimal user disruption.



EMC cloud storage functionality in the virtualized MOSS 2007 environment

EMC Atmos is a cloud-optimized storage platform designed for unstructured content such as office documents and images commonly managed using Microsoft SharePoint. Atmos provides efficient replication and distribution of this content to widely dispersed geographies.

Atmos is designed to scale well horizontally. Additional Atmos segments can be added and removed dynamically without disruption to the existing, operating Atmos storage system. This provides an unprecedented level of durability and high availability for SharePoint users.

When SharePoint content is externalized to Atmos, business growth is no longer encumbered by content growth. In addition, limitations recommended by Microsoft in respect to native SQL Server database storage such as SharePoint database size and document file size are no longer barriers when the content is externalized to Atmos. CPU resources are freed for other tasks or processes that may have traditionally been used for SQL Server overhead and SharePoint data can now be replicated geographically.

Further, large quantities of unstructured data are no longer maintained in SQL Server, which typically uses more expensive storage subsystems such as a fibre-attached storage area network (SAN).

Testing methodology and results

Overview

End-to-end testing of the entire infrastructure was performed to validate the achievable performance levels. Testing was conducted in two key phases:

- **Integration testing—**Validates interoperability and functional aspects of the CLARiiON array, Atmos on premise, and StoragePoint.
- **Scalability and performance testing (externalized versus non-externalized analysis)—**Testing was performed to compare site collections performance that includes:
  - Generating a baseline performance between the two site collections (SQL, Atmos) with three usage profiles (Browse/Search/Modify)
  - Comparing the three methods of user load using Microsoft Visual Studio Team System 2005 (VSTS)

Software used for content creation and client simulation

These tools created the data used in the simulated test environment:

- **KnowledgeLake DocLoader Lite—**Created the content document mix for testing, and populated the SharePoint databases with random user data. Used to provide continuous data population during testing to simulate SharePoint user data growth.
- **Microsoft Visual Studio Team System 2005 (VSTS)—**A platform tool used to emulate client user load while facilitating the design and testing tasks of software teams by improving communication and collaboration.

User and content load profiles used in testing

Heavy user load profile used

During validation, a Microsoft heavy user load profile was used to determine the maximum user count that the Microsoft SharePoint 2007 server farm could sustain while ensuring that average response times remained within acceptable limits.

Microsoft standards state that a heavy user performs 60 requests per hour (or a request every 60 seconds).

Acceptable user response times

The following table lists acceptable limits for Microsoft SharePoint 2007 user operations.

Type of Operation	Example	Acceptable user response time
Common	Browse	< 3 seconds
Common	Search	< 3 seconds
Uncommon	Modify	< 5 seconds

User profiles and response times

Three user profiles (based on a heavy load) were tested to help determine scalability. The following table lists the user profiles and their corresponding response times.

Note
The concurrency rate shown below follows Microsoft's standard guideline for enterprise SharePoint environments.

User activity as % (Browse / Search / Modify)	Concurrency	Average user response time in seconds
80 / 10 / 10	1%	< 3 / < 3 / < 3
70 / 5 / 25	1%	< 3 / < 3 / < 3
50 / 20 / 30	1%	< 3 / < 3 / < 3

Content load profile

The KLDocLoader Lite load profile function evenly distributed the documents across all 30 content databases in the SharePoint farm. The following table summarizes the document mix.

Content type	Number of documents	Doc size (kb)
.doc	891,044	251
.docx	891,044	102
.xlsx	891,044	20
.pptx	891,044	189
.mpp	891,044	235
.jpg	3,691,471	93
.gif	3,691,471	75
.vsd	891,044	471
Total	12,729,206	1,750,010,920

Integration testing

Integration testing

The EMC team performed integration testing to ensure full functionality between the essential components used in the test environment. The integration tests involved the following hardware and software:

- CLARiiON CX4-480 storage array
- Atmos on premise
- Atmos and StoragePoint functionality

CX4-480 storage array

Default testing for storage involved creating LUNs and making them accessible to the servers.

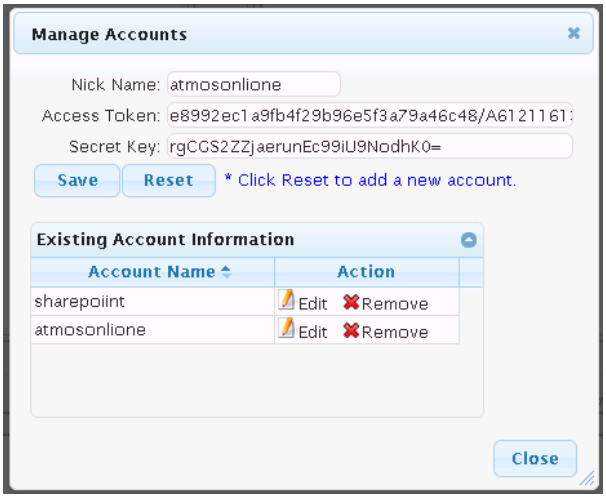
Provisioning Atmos storage

Cloud storage with Atmos isn't provisioned like traditional SAN storage. Using the Atmos GUI, an administrator creates subtenant accounts, and Atmos automatically provisions the storage as required.

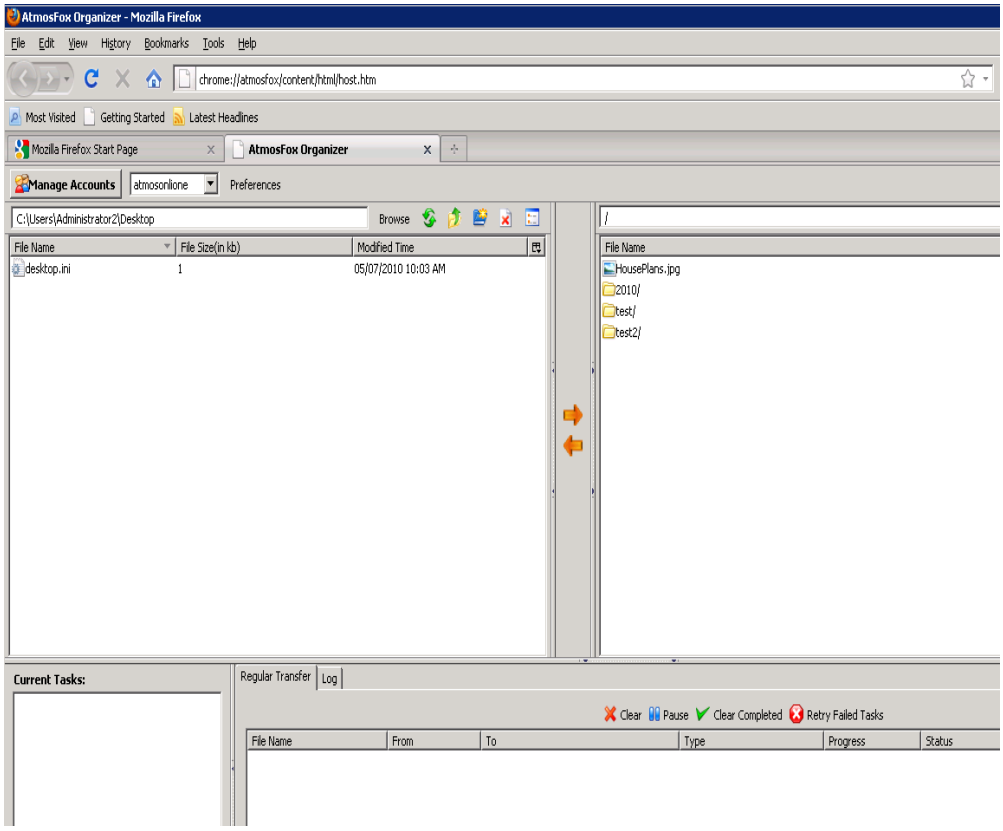
Atmos on premise

Atmos on-premise testing demonstrates how to use Atmos to externalize unstructured data. The AtmosFox plug-in is a helpful debugging tool that is used to test the Atmos setup with the Firefox browser, and is available from http://www.suchisoft.com/atmosfox/AtmosFox.xpi.

The AtmosFox plug-in debugs any connection problems and writes test documents into Atmos. The plug-in is configured as shown in the following image.



When configuration is complete, the AtmosFox plug-in is used to view and debug the Atmos connection setting in the test environment, as shown in the following image.



Atmos and StoragePoint functionality

This validation is reliant on StoragePoint for seamlessly migrating SharePoint objects to Atmos cloud storage. StoragePoint provides several options for simplifying the transition from traditional storage to an external BLOB store. For example, with common implementations of SharePoint in cloud storage customers are presented with only one real option: a hosted model where everything is in the cloud.

By adopting StoragePoint for this function, data centers can host SharePoint servers locally, giving administrators the ability to be able to see and touch the environment. Administrators also have the option of hosting some data locally.

Non-externalized versus externalized analysis

Overview

SharePoint testing targeted two configurations:

- Directly to SQL, Site Collection-1 (non-externalized)
- Through StoragePoint to Atmos on premise, Site Collection-2 (externalized)

Non-externalized test results

Testing resulted in good performance with a roughly 185,000 maximum user capacity for the traditional setup of SharePoint with SQL Server, as indicated in the following table.

Portal usage type	User profile as % (Browse / Search / Modify)[1]	Requests per second[2]	Concurrency	Maximum user capacity[3]	Average user response time in seconds
Heavy read	80 / 10 / 10	30.8	1%	184,800	< 3 / < 3 / < 3
Heavy publishing	70 / 5 / 25	17.5	1%	105,000	< 3 / < 3 / < 3
Heavy collaboration	50 / 20 / 30	15.2	1%	91,200	< 3 / < 3 / < 3

[1] All users were run against a Microsoft heavy user profile (60 requests per hour).
[2] 0% think time was applied to all tests. "0% think time" is the elimination of typical user decision time when browsing, searching, or modifying.
[3] The maximum user capacity is derived from the following formula:
*Second per hour / RPH / Concurrency % * RPS*
Example: 3600 / 60 / 1% * 54 = 324,000
Example: 3600 / 60 / 10% * 54 = 32,400

Externalized test results

During testing, the entire SharePoint farm was externalized (100 percent of data). BLOB externalization completed in 120 hours.

More importantly, Atmos on-premise testing shows that the performance is nearly identical to the traditional setup of SharePoint with SQL as indicated in the following table. Results indicate that relocating BLOBs to an external BLOB store (EBS) shows no impact to the overall user experience.

Portal usage type	User profile as % (Browse / Search / Modify)[4]	Requests per second[5]	Concurrency	Maximum user capacity[6]	Average response time in seconds
Heavy read	80 / 10 / 10	30.1	1%	180,600	< 3 / < 3 / < 3
Heavy publishing	70 / 5 / 25	17.1	1%	102,600	< 3 / < 3 / < 3
Heavy collaboration	50 / 20 / 30	14.8	1%	88,800	< 3 / < 3 / < 3

[4] All users were run against a Microsoft heavy user profile (60 requests per hour).
[5] 0% think time was applied to all tests. "0% think time" is the elimination of typical user decision time when browsing, searching, or modifying.
[6] The maximum user capacity is derived from the following formula:
*Second per hour / RPH / Concurrency % * RPS*
Example: 3600 / 60 / 1% * 54 = 324,000
Example: 3600 / 60 / 10% * 54 = 32,400

Size of databases before and after externalization

The following table illustrates a marked decrease in SQL Server content database size after BLOB externalization completed.

Database	SQL Server DB size (baseline)	SQL Server DB size after BLOB is externalized to Atmos	SQL Server DB reduction ratio
Content_DB_01	106931.67 GB	8997.93 GB	11.88403
Content_DB_02	107871.49 GB	9842.53 GB	10.95973
Content_DB_03	106943.74 GB	8893.73 GB	12.02462
Content_DB_04	106852.48 GB	8869.66 GB	12.04696
Content_DB_05	107139.32 GB	8881.71 GB	12.06292
Content_DB_06	107947.61 GB	9854.5 GB	10.95414
Content_DB_07	107994.39 GB	9876.52 GB	10.93446
Content_DB_08	107507.59 GB	9870.97 GB	10.89129
Content_DB_09	106791.4 GB	8862.66 GB	12.04959
Content_DB_10	111028.51 GB	9916.59 GB	11.196.24
Content_DB_11	108134.7 GB	10040.76 GB	10.76957
Content_DB_12	106830.71 GB	9039.05 GB	11.8188
Content_DB_13	106709.21 GB	9026.59 GB	11.82165
Content_DB_14	107200.81 GB	9053.52 GB	11.84079
Content_DB_15	106889.8 GB	9161.41 GB	11.6674
Content_DB_16	106997 GB	9197.34 GB	11.63347
Content_DB_17	107682.71 GB	10049.69 GB	10.71503
Content_DB_18	107274.59 GB	9053.4 GB	11.84909
Content_DB_19	106885.96 GB	9053.88 GB	11.80554
Content_DB_20	107120.4 GB	9090.39 GB	11.78392
Content_DB_21	106867.73 GB	8898.28 GB	12.00993
Content_DB_22	106687.23 GB	8909.27 GB	11.97486
Content_DB_23	107962.14 GB	9941.3 GB	10.85996
Content_DB_24	110900.64 GB	10936.05 GB	10.14083
Content_DB_25	106878.98 GB	8884.8 GB	12.02942
Content_DB_26	107080.43 GB	8885.79 GB	12.05075

Database	SQL Server DB size (baseline)	SQL Server DB size after BLOB is externalized to Atmos	SQL Server DB reduction ratio
Content_DB_27	111002.72 GB	9909.22 GB	11.20196
Content_DB_28	111422.69 GB	10005.67 GB	11.13595
Content_DB_29	106967.73 GB	8917.62 GB	11.9951
Content_DB_30	107242.15 GB	8930.27 GB	12.00884

Conclusion

Summary

IT planners and storage architects should consider an Atmos cloud storage deployment for their enterprise-level SharePoint environments. By integrating cloud-enabling technologies like Atmos and Metalogix StoragePoint combined with Microsoft Hyper-V for server virtualization IT managers can utilize infrastructure resources and as a result, support an endless capacity for database growth.

This external BLOB storage solution provides a tested methodology designed specifically for an enterprise SharePoint farm without performance limitations, or impact to production within the SharePoint farm. It is a proven model for eliminating the constraints commonly associated with traditional methods of maintaining enterprise-level SharePoint content databases.

Organizations that choose to move forward with cloud-optimized storage enabled by Atmos, StoragePoint, and the virtualization features of Microsoft Hyper-V will:

- Extend their existing SharePoint environment to an enterprise-ready, external cloud service while maintaining overall SharePoint performance, its broad feature set, and user experience
- Realize the value of cloud-optimized infrastructures: faster information access
- Increase operational efficiency while reducing management complexity
- Control the protection, performance, and storage maintenance levels associated with SharePoint content through cloud-optimized protection

Findings

This white paper validates that:

- Atmos enabled by StoragePoint is an ideal BLOB store (EBS) solution that automatically moves unstructured data out of SharePoint content databases leaving only metadata. This reduces the size and number of the content databases.
- External BLOB storage provided by StoragePoint and Atmos produces minimal limitations to user experience.
- The size of the SharePoint databases decreases; subsequently reducing maintenance windows.
- Indexing took 33 percent longer when externalizing SharePoint content databases to Atmos on premise.

Next steps

To learn more about this and other technical validations contact an EMC representative or visit: www.emc.com.

References

White papers

For additional information, see the *EMC Atmos Cloud Optimized Storage for Web Services—Best Practices Planning* white paper.

Product documentation

For additional information, see the product documents listed below.

- *EMC Atmos Release Notes*
- *EMC Atmos Administrator's Guide, Version 1.3.2, Revision 1.3.0A*
- *EMC Atmos Programmer's Guide, Revision 1.3.0A*
- *EMC Atmos System Management API Guide, Revision 1.3.0A*
- *Metalogix StoragePoint Installation and Administration Guide*