Instrumental

HIGH PERFORMANCE INNOVATION

Minnesota Historical Society

Report on Digital Preservation and Cloud Services

Prepared for:

Mary Green Toussaint and Shawn Rounds

Minnesota Historical Society

April 1, 2013

Shawn.Rounds@mnhs.org

Instrumental, Inc.
1450 Energy Park Drive
Suite 375
St. Paul, MN 55108
651-280-4800
http://www.instrumental.com

Contents

Table of Figures

Table of Tables

Notice

Tasking

Instrumental stated in their response that they will help the Minnesota Historical Society (MHS) with its decisions in the comparison of costs and benefits of increasing its existing storage architecture versus purchasing vendor storage services such as cloud storage.

Instrumental believes that all of the areas MHS outlined in the Request for Proposal (RFP) are critical to fully evaluate the cloud market space, and will include additional information based on sources like research presented by Dr. David Rosenthal from Stanford to the Library of Congress during their yearly storage meeting. For example, Dr. Rosenthal stated that many cloud providers do not provide information about reliability, nor will agree to liability or provide information on file fixity.

To better understand the MHS requirements and the market issues, Instrumental gathered requirements such as the following during the initial meeting.

1. Average daily, weekly and monthly retrieval rates (both file counts and data volume)? Most of the cloud storage providers charge not only for archive but retrieval.
2. Peak daily, weekly and monthly retrieval rates (both file counts and data volume)? Some of the cloud storage providers have higher peak rates.
3. What is the timeframe required for a retrieval? Will retrieval needed to be completed in a specific timeframe?
4. Authentication model requirements i.e. who can put data into the archive, who can replace files, who can remove files and who can access files? Do different users have different roles? How many users and does user access need to be tracked?
5. What is the incoming and outgoing bandwidth to MHS-This is critical to understand to determine the maximum potential data access.
6. Will MHS accept LTFS on LTO5+ for large data volume restoration? This is likely much less expensive for large amounts of data needed at MHS.
7. What is the technical expertise of the MHS employees overseeing the digital archive? Some cloud storage providers can be more turnkey while others require highly technical staff to operate.
8. What are the plans for future growth?

In the final report and briefing, Instrumental will address and help MHS understand all of the factors outlined in the RFP including background on content related issues to consider, cloud services vendors to consider, costs, best practices, as well as other issues pertinent to the MHS data preservation mission.

Overview

Today many large data sites are considering complete outsourcing of their archives to cloud vendors, but there are many things to consider when moving to an outsourced archive provider. Some of the areas that the Minnesota Historical Society (MHS) have expressed concern with are extremely well founded such as silent data corruption, and Instrumental's CEO/CTO has published and spoken on this topic for the last 5 years.

Instrumental recently supported customers with similar larger archives and similar questions and cost / benefit analysis requirements such as the Library of Congress, DoD High Performance Computing Modernization Program, and NOAA Comprehensive Large Array Stewardship System and NOAA Geophysical Fluid Dynamics Laboratory.

To clarify terms used in this report, the terms backup and archive are defined below so that there is no confusion when these terms are referenced elsewhere in the report.

- Backup – This involves copying and saving data so it may be used to restore the original after a data loss event, whether by data deletion or corruption. It can also be used to access older data that is no longer stored on disk.
- Archive – This is a file that may contain one or more other files for data retention. Most archive formats are also capable of storing folder structures in order to reconstruct the file/folder relationship when files are restored.

Background on Content Related Issues

The following sections describe some of the content related issues for the MHS preservation activities. As a part of understanding the content, the current information on collections maintained by MHS is shown in the table below.

Content Type/Collection	1/2013 Approximate Quantity (TB)	Projected Additions	Projected Annual Increase (TB)	Anticipated Retrieval Needs	Exposure Risk/Privacy Concerns	Value	Notes
Birth Records	8.0	None	1.2	None	HIgh risk profile, confidential info included	High, accessioned collection, not replaceable, ecommerce, master set for state	Should have at least 3 backups, with one off-site
Accessioned Collections (aggregate, excluding Birth)	1.0	None to individual collections	0.1	None	Low	High, accessioned collections, not replaceable	Should have at least 2 backups, with one off-site. New collections will be added over time.
Unaccessioned, High-Value Collections (aggregate)	23.0	File batches at unscheduled intervals to individual collections	3.4	None	Low	High, fragile originals, ecommerce	Should have at least 2 backups, with one off-site. New collections added over time.
Unaccessioned, Low-Value Collections (aggregate)	1.0	File batches at unscheduled intervals to individual collections	0.1	None	Low	Low, replaceable	Should have at least 1 backup. New collections added over time.
NDNP MN	22.0	None after FY2014	0.6	Once loaded into NMS, NONE	Low - pre1923 out of copyright	TIFFs are not replaceable without redigitizing from microfilm. Tifs are backed up at LC, but not available for download.	Should have at least 2 backups, with one off-site. Remaining files are backed up at LC and available for download.
NDNP North Dakota	7.0	In-progress	2.8				Project expected to complete in FY18. Data returns to ND. First 6 microfilm digitized are from MHS master negatives.
NDNP Iowa	0.0	In-progress	3.5				Project expected to complete in FY19. Data returns to IA.
SELCO	20.0	In-progress, may be on-going	3.5	Once loaded into NMS, None	Low - pre1923 out of copyright	High value, need to redigitize if lost	Should have at least 2 backups, with one off-site
MHS Newspaper Collection	0.0	In-progress	4.0	Once loaded into NMS, None	Low - pre1923 out of copyright	High value, need to redigitize if lost	Should have at least 2 backups, with one off-site
MNA Digital Newspapers	0.0	In-progress	8.0	Once loaded into NMS, None	High- still in puublisher copyright	MN statuate mandates that MHS preserves copies of legal newspapers, born digital	Should have at least 3 backups, with one off-site
Oral History	2.3	None	0.0	Once loaded into CMS, None	Some have restrictions.	High, not replaceable, current OH are born digital	Should have at least 3 backups, with one off-site
Swedish-American newspapers	13.0	Upcoming	0.0	Once loaded into NMS, None	Low - pre1923 out of copyright	High value, need to redigitize if lost	Should have at least 3 backups, with one off-site
CMS multimedia	3.6	Ongoing additions of files as new files are added to the CMS.	2.1	TIFF files retrieved to fulfill photo orders	Some have restrictions.	High value. Some would need to be re-digitized if lost. Others are born digital and could not be replaced.	Should have at least 3 backups, with one off-site. WAV and other media master files may also occasionally be needed to fulfill orders or for MHS use

Table 1 – MHS Collections Effective January, 2013

Ownership of Files

As can be seen from the table above, some collections in the archive belong to third parties and the data will eventually be returned to the owners. These collections must maintain their ownership rights regardless of where the data is stored (locally or in the cloud). In addition, some of the collections contain proprietary and copyrighted data for which MHS is mandated to maintain control. Most cloud providers ensure standard user rights of ownership but some such as Google include language that gives them more control over user data.

Contractual Obligations

As already noted, MHS has proprietary and copyrighted information for which it has contractual obligations to protect. In addition, there are contractual requirements to return some data collections to its owners upon completion of the collections. For this reason, maintaining the integrity and availability of the data in the collections has significant legal ramifications.

Storage Type Reliability and Integrity

The hard error rate, which defines how many bits of data can be read before a read fails, for enterprise tape storage is at least 2 orders of magnitude better than that for even enterprise disk storage as shown in the table below.

Device	Hard Error Rate (1 bit in error in this number of bits moved)	Equivalent in Bytes	PiB Equivalent Data Moved Before Error
SATA Consumer	1.0E+14	1.25E+13	0.01
SATA Enterprise	1.0E+15	1.25E+14	0.11
Enterprise SAS/FC	1.0E+16	1.25E+15	1.11
LTO and Some Enterprise SSDs	1.0E+17	1.25E+16	11.10
Enterprise Tape	1.0E+19	1.25E+18	1110.22

Table 2 – Disk and Tape Hard Error Rates

Below is what these hard error rates translate to in terms of failure rates with the devices running at their rated speeds.

	Number of Devices				
	1	10	50	100	200
Device Type	Hours to Reach Hard Error Rate at Sustained Data Rate				
SATA Consumer	50.9	5.1	1.0	0.5	0.3
SATA Enterprise	301.0	30.1	6.0	3.0	1.5
Enterprise SAS/FC 3.5 inch	2,759.5	276.0	55.2	27.6	13.8
Enterprise SAS/FC 2.5 inch	1,965.2	196.5	39.3	19.7	9.8
LTO-5	23,652.6	2,365.3	473.1	236.5	118.3
LTO-6	20,696.1	2,069.6	413.9	207.0	103.5
Some Enterprise SAS SSDs	7,884.2	788.4	157.7	78.8	39.4
Enterprise Tape	1,379,737.1	137,973.7	27,594.7	13,797.4	6,898.7

Table 3 – Failure Impact of Hard Error Rates

Based on the industry defined hard error rates, options with a tape storage component will have a higher level of reliability than disk or SSD. It should also be noted that a single bit in error on a disk drive translates to a full sector error (512 bytes or 4096 bytes for some newer devices) while with tape a bit in error results in only 1 bit of data being lost.

Since no cloud vendors provide any data integrity guarantees, a requirement is that there be checksums generated before the data is moved to the cloud to ensure that data is correctly stored in the cloud and that the data retrieved from the cloud is identical to what was stored. These locally maintained checksums must be periodically validated at least until MHS is confident that the cloud data integrity is assured. This may be a number of years. In the Summary and Analysis section later in this document, a proposed framework for verifying data integrity is indicated. Some of the cloud storage providers have this capability while others do not. Most cloud providers have similar standard reliability capabilities by having multiple data copies.

Disaster Recovery

A cloud storage option can provide an excellent disaster recovery mechanism if the amount of the data to be recovered is not too large or if high speed networking is available. Other options for disaster recovery include the shipment of disks or tapes if compatible equipment is available at the customer site. Most cloud vendors have additional charges for this service.

Retention and Portability

One area of concern is the retention and portability policies of the cloud vendors. Some cloud vendors make it difficult, time consuming and expensive to move data from their cloud service to that of another vendor. There is also the issue of ensuring that all copies of the data have been removed once an agreement with a vendor has been terminated; in particular, note the language in the Google contract as detailed in the Google and Google DRA section below.

Availability

Some of the vendors provide an availability (uptime) guarantee; these are fairly standard and range from 99.9% for AWS, 99.98% for VISI and to 99.99% for Permivault. Others such as Google and IBM SmartCloud do not provide any availability guarantees, but often make marketing availability claims that they will not provide in a contract. Still others, such as DuraCloud and Tessella Preservica, have an assumed availability based on the cloud providers that they utilize. MHS should look at issues such as the Microsoft Azure down time in late February to understand that cloud provider claims of availability sometimes cannot be met.[1]

Breaches

The cloud vendors have extensive documentation on security features; an example is the whitepaper on AWS security.[2] However, there is limited or no information on how data

[1] Microsoft Secure Azure Storage Goes down Worldwide; http://www.theregister.co.uk/2013/02/22/azure_problem_that_should_never_happen_ever/
[2] Amazon Web Services: Overview of Security Processes; http://aws.amazon.com/security

breaches are handled. This may be intentional to provide as little information as possible to those who are attempting to breach the cloud provider facilities and gain access to the data stored there. With cloud service providers aggregating access to many victims' data into a single point of entry and as their services become more popular, they will increasingly become the focus of attacks, experts say. Encryption of data is therefore a necessity in this threatening environment; to support this, MHS will need to develop a key management plan for their data.

Scalability

For most, but not all of the cloud providers surveyed, scalability does not appear to be an issue for the size of the collections that MHS currently has and most offer scalability to 1000s of TiB or more. The one exception is VISI, which currently does not offer the required level of storage in a manner that is cost competitive.

Data Management and Preservation Functionality

The various cloud providers have differing data management and preservation functions ranging from little or none for providers such as AWS and Google to significant functionality for products such Tessella Preservica and DuraCloud. For example, Tessella Preservica allows users to manage and store data using metadata and data tags; the metadata can be read and recorded in a database according to marketing claims. Preservica also offers web-based interfaces for browsing a collection or searching for content. Regardless of the provider, MHS needs tools to edit and dispose of information in a controlled manner; these include editing of metadata, disposition to third party locations, hard and soft deletion and approval cycles.

Preservation is provided in Preservica by background integrity checks, which look for missing files and file corruption, and it is also claimed that it includes a set of tools to migrate files from obsolete formats as technology changes, but this will be difficult to accomplish if the file names and data are encrypted. As another preservation example, DuraCloud has several ways to verify MD5 checksums; users can manually run the verifications or DuraCloud can run them automatically at defined periods. Another option is that DuraCloud can provide automatic checksums upon file retrieval.

Other Issues

Technology obsolescence is an issue that can affect both content and the metadata associated with the content; archiving systems need to provide mechanisms to monitor and transform content as needed to protect against obsolescence. Tessella Preservica currently claims to provide this capability by offering tools to migrate files from obsolete formats as technology changes, while DuraCloud plans to expand this capability beyond images in the near future. However, this may not be feasible with the challenges of encrypted file names and data.

Cloud Services Vendors

The following sections provide details of the cloud services vendors.

Amazon S3 and Glacier

Amazon S3 was originally intended as a simplified web storage space to store and retrieve data at anytime from anywhere with an initial emphasis on developers. Over time, this emphasis shifted to reach larger customer bases with large cloud storage options. Amazon S3 reaches this goal by offering storage options over 5000 TiB of data[3]. Amazon also offers higher availability and data uploads by offering to employ AWS Import/Export. Amazon makes marketing claims (these are different than contractual commitments) that, with an 80% commitment from the network, 60 TiB of data can be transferred either way using a 1000 Mb/s connection in less than one week. For data integrity, Amazon S3 offers redundant storage in multiple facilities or on multiple devices while calculating checksums on all network traffic to detect corruption of data packets when storing or retrieving data. S3 marketing claims that it performs regular systematic checks and claims to be built to be automatically "self-healing"; however, there is nothing backing this in the service level agreement (SLA).

For security, S3 offers a variety of authentication mechanisms, rights-granting options, and encryption options. Only data owners are allowed access to the data while AWS's identity and access management and access control lists allow owners to limit access. For encryption, SSL encryption endpoints using HTTPS protocol for uploads and downloads is offered while S3 also offers a no-charge option for owners to manage their own encryption keys with client-side encryption. Server encryption is AES-256. S3 also offers auditing and access log recording options. This is confirmed to a degree by Amazon S3's PCI and HIPAA compliant status. S3 also allows for multiple management options including data lifecycle management that allows for managing capacity, deletions and transfers of data. Actual contractual obligations are based on the S3 SLA and Amazon Web Services terms of service. The S3 service level agreement commits to at least 99.9% uptime for availability and offers 10% discounts if uptime drops below 99.9% and 25% discounts if uptime drops below 99%. This is done on monthly bills that Amazon sends and is calculated based on the number of times Amazon S3 returns "InternalError" or "ServiceUnavailable" responses to attempts to access S3. However, under the AWS customer agreement, it states that the user/owner of data is solely responsible for maintenance and security of content. There is no guarantee in the terms of service that content will be uninterrupted, error free or not lost; Amazon is not liable for any loss of data.

One unique aspect of S3 is that data can easily be transferred between it and Amazon Glacier. Glacier was created as a service that offered long-term storage and retrieval for large data sets at a low cost[4]. It offers storage spaces over 5000 TiB (like S3) with separate archives of up to 40 TiB. Availability is much more limited, however as Amazon states, retrieval of data can take between 3 to 5 hours for individual files or parts of archives; data is only made available for 24 hours after retrieval. Integrity marketing claims are similar to S3, but Glacier has the added restriction that archives can be uploaded, read, or deleted, but cannot be edited or overwritten. After an archive is uploaded to a vault, its content or its description cannot be updated. The only way the archive content or its description can be updated is by deleting the original archive and

[3]Amazon Simple Storage Service; http://aws.amazon.com/s3/
[4] Amazon Glacier; http://aws.amazon.com/glacier/

uploading another version of the archive. Note that each time a customer uploads an archive, Amazon Glacier returns a unique archive ID. Amazon states that not allowing an archive to be updated or overwritten prevents outside or accidental tampering while limiting damage to "bit-rot." This statement is only partially correct from a technical standpoint. Making files read-only limits tampering but has no impact on bit-rot. It will increase customer costs if the customer needs to update a file since there will be additional upload costs.

Security is similar to S3 as all data is encrypted on the server side with AES-256 encryption and Amazon manages and handles key management and protection. However, customers can manage their own encryption keys if they encrypt data prior to uploading it to Glacier. By default, only the owner can access data and controls access by using AWS identity and access management to define access to each vault separately (vaults are used to separate different types of archives or collections). While compliances are never guaranteed, Glacier's interactivity with S3 means that it should also be HIPAA compliant since S3 is HIPAA compliant or otherwise S3's HIPAA compliance for customers in medical fields would be compromised, but this is implied not stated. For control, the owner creates and manages archives and the vaults in which the archives are stored. For contractual obligations, Glacier follows the same route as S3; however, there is no guarantee that uptime rules for S3 apply to Glacier. For backup and recovery, Glacier does offer the ability to replace or supply magnetic tapes, but types, formats, and costs are not specified on the web site.

Google and Google DRA

Google's Cloud Storage service began as, and still is, focused on supporting Google developers and application developers by allowing a space to store data. This cloud storage provider can store over 5000 TiB of data and offers a variety of availability options based on user needs and budgets. As far as integrity and checksums, Google makes very few claims beyond generic marketing and advertising and actually warns that data integrity could be an issue for those looking for long-term storage for preservation. In addition, security measures are limited. While most cloud providers pride themselves in security measures that they take or security compliances (such as HIPAA in order to market to potential clients in the medical field) that they have, Google actually downplays these features and limits security measures to a Google account and Oauth 2.0 authentication mechanisms[5]. For management, a Google API console is required and access management is an "all-or-nothing" format where other users either have full access (including the ability to write to data) or no access at all. For advanced management features, a command-line tool (*gsutil*) is required.

Durable Reduced Availability (DRA) storage enables customers to store data at lower cost, with the tradeoff of lower availability than standard Google Cloud Storage. It has the same durability as standard Google cloud storage and is appropriate for applications that are particularly cost-sensitive, or for which some unavailability is acceptable. Examples include:

- Data backup – This is a case where high durability is critical, but the highest availability is not required.

[5] Google Cloud Storage; https://cloud.google.com/products/cloud-storage

- Batch jobs – Batch jobs can recover from unavailable data, for example, by keeping track of the last object that was processed and resuming from that point upon restarting.

The actual terms of service agreement is just as alarming. It states that "[by] submitting…Data on or through [Google Services], the Customer gives Google a worldwide…limited license to reproduce, adapt, modify, translate, publish, publicly perform, publicly display and distribute any Customer Data for the sole purpose of enabling Google to provide and improve Services." While this may not threaten MHS (as none of their data would realistically improve services), the fact that Google can access and reproduce a user's data without warning is contrary to most cloud providers, who usually deny themselves any access to a user's data. Google Cloud Storage appears to remain focused on short-term storage for Google application developers and no other potential data storage clients. Therefore, any data put into Google's environment will need to be encrypted by MHS before placing into the Google system.

Tessella Preservica

Preservica claims to have a specific focus on data archiving and long-term preservation as its niches in cloud storage[6]. For its base architecture and storage, Preservica uses AWS, so its overall storage capacity for a user should be over 5000 TiB like Amazon's cloud storage services (S3 and Glacier). Tessella makes no specific marketing claims on availability other than that it can be fast, based on size of the package and workflow. Preservica states that they offer high data integrity through the creation of multiple copies stored in multiple data centers that are checked against one another and uses a combination of content checksums, either at customized intervals or upon retrieval of data, and cyclic redundancy to regularly check for data corruption. Security runs on the AWS-hosted systems and is based on user access controls. No third parties, including Amazon, are granted access to data and users can control access to both data and metadata. Tessella states that their encryption options are based on Amazon's encryption options so AES-256 encryption is available.

Preservica highlights that users can manage and store data using metadata and data tags; marketing also claims that metadata can be read and recorded in relational databases while it offers web-based interfaces for browsing a collection or searching for content. Tessella has done a proof of concept for emulation to recreate older hardware environments virtually to access old information, but this capability is not currently part of Preservica. While Preservica is primarily disk-focused, it does offer magnetic tapes as a possibility for either part of the regular storage option or as a form of secondary media backup to be used for retrieval.

In a call with Tessella on 2/20, the Preservica software was described; it utilizes the Safety Deposit Box (SDB) software running on AWS. Preservica is a somewhat pared down version of the SDB software that is run at customer locations. SDB has plug-ins for various storage options such as NAS, SAN, or tapes, but these options would not be available with Preservica on AWS. This cloud version of SDB used by Preservica is at a lower cost than the standard SDB and offers a lower barrier of entry. Preservica offers more in the way of preservation features than the other

[6] Preservica – Digital Preservation as a Service; http://www.digital-preservation.com/solution/preservica/

vendors, but at a higher cost. Data preservation is claimed to be guaranteed by Preservica's automatic regular background integrity checks, looking for missing files and file corruption. Preservica also includes a set of tools to migrate files away from obsolete formats as technology changes. One possibility is to use Preservica only for the most high value collections at MHS and use another cloud service for the less valuable collections. However, this approach would increase complexity and the management of different archival solutions would be more time-consuming.

VISI

VISI's ReliaCloud aims to provide "high-availability, scalable performance, and comprehensive security" for clients considering cloud storage[7]. While no specific information is provided on the amount of data ReliaCloud can handle, VISI facilities can handle 10 Gb/sec Ethernet (GbE) networks; it claims its architecture will have future support for 40 GbE. However, representatives stated that the current nature of their arrays makes over 150 TiB of data difficult to handle and cost prohibitive, but that upgrades to an EMC Isilon NAS solution and systems over the next year could be able to fulfill the MHS data storage needs. This information was provided by VISI representatives in a meeting on February 14, 2013. Aside from claims on its Ethernet network, availability is based on bandwidth performance. VISI provides dedicated 1 GbE Internet connections for downloading and uploading data, so the speed of upload is based primarily on the connections MHS can use.

VISI protects data integrity through both physical and electronic means. The company provides a 100% SLA-backed guarantee for availability, but not for integrity. Their facilities Uptime Institute Tier III certified (guaranteeing 99.98% uptime), are guarded and monitored 24 hours a day, 7 days a week and incorporate advanced fire detection and suppression equipment as well as redundant HVAC equipment to preserve the environment of data storage rooms. Facilities also utilize redundant uninterrupted power systems, backup generators, redundant power paths and supplier agreements to ensure availability. VISI relies on this facility security and uptime maintenance, including disk backups, as a means to maintain data integrity. However, they could provide no information regarding self-healing or checksum abilities, and technologists with VISI did not understand questions addressing checksums or bit-rot.

In addition to these physical safeguards, systems use redundant CPUs, memory, and connectivity on two directors to provide electronic backups of data. Multiple data centers are located in Eden Prairie, Minnesota; Des Moines, Iowa; and Madison, Wisconsin; they are linked by dedicated private 10 GbE lines that offer the possibility of replication and multiple copies in multiple locations. In addition to several physical security measures, including storage areas secured with biometric readers, electronic security is a priority. Firewalls, third party encryption from Vormetric, and data owner access controls in a least-privilege environment with customizable data partitions are provided in compliance with HIPAA and PCI DSS requirements.

[7] Cloud Services; http://www.visi.com/services/cloud-services/

VISI offers local storage in the Saint Paul and Eden Prairie areas. Another important note is that their acceptable use policy includes a clause which allows users to notify security of improper accesses, copying, or writing to data, and VISI will make efforts to prevent further loss or replace losses with backups if possible, but replacement is not guaranteed. Data management can be done through either an FTP-based or web server-based interface to browse and upload data files. VISI ReliaCloud works primarily with disks and currently has no systems that integrate magnetic tapes as a storage media, but are willing to work with MHS to architect a solution that could include tape preservation or recovery.

SDSC's Cloud Storage

The San Diego Supercomputing Center's Cloud Storage offering was originally designed for academic resources and research environments. They provide storage options, hypothetically, greater than 100 PiB for large scientific data flows and projects[8]. SDSC marketing claims that proper load-balancing and automated failover ensure continuous access. Marketing also states that 10 GbE switching provides sustained read rates of up to 8 GiB/sec on SDSC's end of the connection and they offer the possibility of streaming data. SDSC also claims that two copies will be stored on two separate servers at all times and that error-checking is continuously run to preserve data integrity and accuracy. Additional offsite storage for disaster recovery is housed at UC Berkeley; however, both sites are on major fault lines and at risk from earthquakes. For security, AES 256-bit server-side encryption is offered and user access controls are provided to ensure HIPAA compliance for medical research. SDSC uses the Rocks clustering tool kit for management, or users can choose to use interfaces that work with Rackspace or S3 through OpenStack "swift" object software. The SDSC also offers no bandwidth charges but they plan to provide on demand dual-site storage options within the next year. The system runs 100% on disks and does not offer options for magnetic tape storage and SDSC does not plan to have a tape backup. This is a significant issue given disk head crash potential during earthquakes and having no copies at rest.

DuraCloud

DuraSpace's DuraCloud was designed for storage and preservation for research centers and cultural heritage organizations through multiple service providers[9]. These providers include AWS, Rackspace and SDSC Cloud Storage; available in April, Amazon Glacier will be another lower cost option. Since it uses other providers, DuraCloud's maximum space is over 5000 TiB of data that operates on a "buy-as-you-store" system where space is purchased as needed. DuraCloud marketing claims that it can provide rapid availability even with streaming for customers, which is provided as an extra cost option, and it can also store and rapidly deliver images and videos. However, without any specific data and its reliance on other cloud providers, the speed of availability will significantly vary depending on configurations, plans and conditions. To address data integrity and preservation, several ways to verify checksums are available either by allowing users to manually run checksum "checkups", DuraCloud can provide automated checksum services or automatic checksums can be provided upon file

[8] SDSC Cloud Environment Overview; https://cloud.sdsc.edu/hp/system.php
[9] DuraCloud Overview; http://duracloud.org/overview

retrieval. They offer a health checkup service for spaces larger than 1 TiB, which uses an Amazon Hadoop cluster to determine checksums for all content items in any particular space.

With connections to multiple cloud providers, DuraCloud offers multiple backups with multiple providers in different locations. Security focuses on a variety of encryption options (including options from providers like Amazon or SDSC) and authentication using REST POST methods and the DuraStore REST API in HTTPS. However, since third parties provide the actual storage, DuraCloud cannot provide assurances related to security of data, according to the terms of service. This third party storage system means that compliances like HIPAA can only come through compliant providers such as Amazon or SDSC since data is ultimately stored in these locations and DuraCloud does not have a privacy policy of its own. Management of data can be done through a single dashboard for all storage, even when using multiple cloud providers. DuraCloud claims it can handle all transactions and negotiations with providers, even allowing for movement of data into, out of and between providers without extra cost. As mentioned above, terms of service point out that reliability, security and availability are actually dependent on the capabilities of cloud providers and not DuraCloud.

IBM SmartCloud

Originally designed as an extension of in-house cloud systems for clients using IBM software and hardware, SmartCloud has been extended to include data storage services[10]. However, it still relies to a certain degree upon clients' use of IBM technology and systems for interaction and added analytic functions. Long-term preservation is a growing priority; with 850 TiB of data per system administrator, IBM marketing claims that it can provide nearly unlimited capacity. IBM also markets that its SmartCloud storage services can provide recovery times of minutes depending on the amount of data, although no specific timeframes or data set sizes are provided for more accurate estimates. Longer times of at least hours are more realistic for larger data sets.

For data integrity, IBM marketing claims that it has the ability to "self-heal," but provides no details on what this capability actually entails. Additional preservation and integrity safety comes with the ability to store copies of data at multiple locations. Customer access management control and advanced encryption techniques with additional data de-duplication provide the backbone of security for data stored on their systems. Here again, there are few specific details provided. Given the claims on security, HIPAA compliance is a possibility, but it is not guaranteed by IBM, and IBM provided no details on this. To manage and sort data, customers need to use IBM's software, management tools, and analytic tools. This requires the installation and use of specialized IBM software. In addition, IBM's cloud services are not well integrated, risking the creation of service "silos" that could limit flexibility and communication between services. The system can use LTFS and LTO 5/6 tape media for storage, ingest and preservation.

In a call with IBM on 2/18, one of their cloud architects was unable to answer a number of the critical questions posed to them such as:

- Is forward error correction available in IBM's cloud storage offering?

[10] IBM SmartCloud Managed Backup; http://www-935.ibm.com/services/us/igs/cloud-data-backup/details.html

- Is bit error level correction available?
- What is the reliability guarantee (especially focused on storage and ensuring there are no bit errors in the archived historical documents is key)
- Upload / migration - What are the costs to initially upload the data and to possibly migrate it if they move to another provider in the future?

One of their senior technical leads on the SmartCloud for the Enterprise offering was to provide answers to these questions, but no response was received from him as of the delivery date of this report. It appears that their cloud services, backup services, and preservation services all are separate from each other and dealing with the IBM bureaucracy is difficult.

Others

The following sections will highlight other potential vendors not included in the list provided by MHS.

Permivault

FujiFilm's Permivault was created as a long-term file bank for unstructured files to ensure safe and secure offsite protection and second copies of files to aid in disaster recovery. Designed for large amounts of image or video storage and through the use of magnetic tape media, Permivault marketing claims to offer nearly unlimited storage space[11]. Availability of data is dependent upon the use of an onsite StrongBox storage unit from Crossroads and on tapes as well as bandwidth of the customer network. The StrongBox storage unit is compatible with LTO-5 tape libraries from Dell, HP, IBM, Overland, Quantum and SpectraLogic. A complete list of the currently tested and qualified tape libraries and software is available from FujiFilm. Appendix B - Permivault Compatible Tape Drives and Software has the current list of tape libraries and software, but this list is likely to change and FujiFilm should be consulted for the latest information. If the customer chooses not to have an on-site StrongBox system, there is client software that provides access to the remote StrongBox system and annual maintenance charges apply for the client software.

Integrity of data is based on custom data protection plans where Permivault checks data on tapes at the customer's request and for recovery. Security options vary, but data can be encrypted through the chain of custody for both inbound and outbound encryption up to 2048-bit encryption. The customer also is able to manage access controls for further security. This is supported by both SSAE 16 and HIPAA compliance. Additional features include a marketing claim of 99.99% uptime and a guarantee of three copies of data on LTO-5 tape backups, with two in active storage and another copy stored at an off-site vault 10 miles away from the data center. In addition, the option to use LTFS and import and export content that can be read by any LTFS-supported tape system is available.

[11] Permivault Solutions: Enterprise-class, Cloud-based Data Protection; http://www.permivault.com/solutions/

Code 42 CrashPlan PRO

CrashPlan PRO combines onsite, offsite and cloud backup in an easy to use, enterprise strength solution.[12] It is cross-platform, client / server backup software that runs on customer hardware. CrashPlan PRO software can also backup to hosted online backup services for offsite, mission critical backup and requires little setup. Each of their data centers complies with the strictest security standards and conforms to industry best practices for power redundancy, cooling and protection symptoms. CrashPlan PRO uses 448-bit Blowfish encryption for files before transferring them and the files are transferred to their servers using AES-128 protocol.

While engineered for laptops, CrashPlan PRO runs in the background on any computer with minimal impact on user applications. It protects and restores customer data with enterprise-capable backup technology. In addition to scheduled backups, the product provides continuous, real-time backup protection. CrashPlan PRO detects data changes then waits for the changes to stabilize. Once the changes stabilize, it only transmits newly created data; if data is merely re-organized, no sending of data is required. This byte pattern differential approach conserves considerable bandwidth and reduces load on users' computers when compared to traditional backup differentials.

A wide variety of clients are supported by CrashPlan PRO, and it provides the ability to limit the amount of CPU and bandwidth being allocated to backup. Once the initial backup is completed, there is typically only a trickle of data sent over the Internet to reflect the changes that have occurred in the data. Marketing claims that their compression and de-duplication are very efficient, which further reduce the transfer requirements. In addition to the backup capabilities, it provides:

- User-initiated restores – Files can be restored quickly with minimal technical expertise
- Universal networking – Works over wireless, wired and cellular networks
- Local data encryption – Secures information before transmission
- Cross-platform support – Support for Windows, Macs, Linux and Solaris

Recommendations by Professional Organizations

A number of professional organizations have made recommendations about what to ask from cloud storage providers. As Instrumental is involved with the NDIIPP, we are aware of the evolution of some of the suggestions, but at this point detailed recommendations about all of the areas to consider have not been fully vetted. Other professional organizations such as SNIA, American Library Association, and government agencies, such as the National Archives and NOAA, have made limited recommendations. Instrumental has researched a number of the recommendations and believes that the only comprehensive set is from LOCKSS.[13] The problem with most recommendations is that they do not specifically address all of the end-to-end problems. Some recommendations make comments like storing 3 copies of data and talk about

[12] CrashPlan PRO; http://www.crashplan.com/business/
[13] Preservation Principles; http://www.lockss.org/about/principles/

using DVDs, but they fail to understand the poor reliability of DVDs. The technology issues are complex as are the frameworks to ensure end-to-end integrity and the security of the information.

Cloud Storage Activities Being Used or Tested

The following sections highlight some of the activities taking place with some of these products in other locations.

National Association of State Chief Information Officers

The cloud is one of many IT priorities for the coming years, and through surveys, NASCIO knows that cloud level adoption is up. They published a four part series called Capitals in the Cloud that covers business objectives, governance, acquisition strategy, jurisdictional issues, security and privacy, policy and legal issues, exit strategies and more. The 2012 State CIO Survey Report: Advancing the C4 Agenda – Balancing Legacy and Innovation results were published in October 2012 and it has a section on cloud computing. When looking at cloud services, the report found that cloud adoption is up, others are developing cloud services and nobody has rejected the cloud.

- 15% are highly invested
- 56% have some adoption
- 19% are developing services
- 48% of respondents plan to use the cloud for storage
- 44% for disaster recover
- 31% for digital archives/electronic records

National Archives and Records Administration

NARA is very focused on electronic records management and how good practices are needed to move forward. They will be producing a report on the federal cloud initiatives. People are concerned with the cost models of cloud services. How do these compare with paper vs. electronic?

Other Issues or Data to Consider

There is little preservation functionality and limited integrity checking present in a number of the package such as Amazon S3 or Glacier or in Google Cloud Storage. The data integrity checking capabilities of CrashPlan Pro and VISI are somewhat better. The Permivault product has good data integrity capability but has limited preservation functionality since this is a new product and has not addressed this need. The Tessella Preservica product is believed to have the most preservation functionality as well as good data integrity control based on their marketing claims, but at a significantly higher price. MHS must consider what capabilities are essential to them and how much they are willing to spend to provide these capabilities.

Staffing Requirements and Definitions

With the majority of the cloud provider offerings, it is believed that the staffing requirements will be fairly minimal once the required data has been ingested into the cloud provider's systems. Probably a part-time system administrator will be adequate to check on the health of the data stored with the cloud provider through performing periodic checksums of the data and

comparing those to the checksums previously saved. Assuming an on-site copy of all data collections is maintained, the incremental work for the data stored in the cloud would not be that great unless there are problems encountered. The option of a remotely monitored Strongbox system as proposed by FujiFilm will further decrease this workload if MHS wishes to exercise this option.

Costs by Vendor Product

Based on the projected annual growth rate of 25% for the MHS collections and starting with 130 TiB in the first year, an estimated cost for the next 3 years was determined based on the current pricing models. While it is likely that the pricing models for a number of the vendors will change, it is difficult to estimate what these changes will be. Because of sensitivity of some vendors to the pricing provided, no actual costs are provided, but rather a pricing range of low, medium, or high is shown in Table 4 - Comparison of Capabilities and Pricing.

A note of caution on pricing is that hidden away in the small print of many cloud service contracts are a range of charges that customers may not be aware of until it is too late, that is, when very large bills for data processing or data storage are charged.[14] Some examples of these practices are:

- Cloud pricing will often incur general usage costs, i.e., pay per use, but a supplier may insist on a minimum charge of 2 hours, for example, even though the customer only requires 20 minutes of compute power.
- Similarly, partial hours are also often rounded-up for charging back. For example, 2 hours and 5 minutes of usage would be charged at 3 hours.
- Users may also find that attractive initial rates apply only up to a certain level of use, and beyond that premium rates kick in, escalating the overall cost.
- Often missed by users in the small print are charges for bandwidth in / out, i.e. the transfer of data. Even if spotted, this is a cost that may be difficult to estimate and even harder to control. So users should be very aware of these charges.
- Suppliers can charge for the number of users accessing the resource – with costs increasing in stages, e.g. 5 users or 10 users – similar to the traditional software licensing model with which most users are familiar.

Summary and Analysis

Instrumental believes MHS needs to consider the following characteristics of a cloud archive when comparing to local management:

1. What is the underlying storage technology which determines the overall reliability of the archive? The hard error rates of disk and tape are not going to change.
2. Given the available service providers and the fact that new providers are coming to market monthly, what customer bases are the cloud providers actually attempting to

[14] Warning to Cloud Adopters: Check the Fine Print; http://www.hpcinthecloud.com/hpccloud/2013-02-25/warning_to_cloud_adopters:_check_the_fine_print.html

serve? Adding a framework for data integrity on top of a cloud provider attempting to serve a different market will not be likely to meet overall requirements. Cloud storage providers are often marketing to application developers which may or may not meet the needs for MHS. MHS needs data integrity and security on a per file basis. MHS should look for providers that are targeting long-term preservation of sensitive and/or historical files.

3. What are the security policies, procedures and certifications in place? Adding security features to a cloud provider with an inherently insecure environment will probably not meet the MHS legal requirements.
4. What is the data validation and integrity framework? MHS needs to have a framework in which integrity can be validated both inside and outside of the MHS cloud archive. MHS must be able to periodically check the integrity of stored data against original copies in addition to using vendor-provided checksums in the cloud environment to ensure integrity of the data after transfer (see the detailed explanation below).

The diagram below shows the framework that Instrumental believes is needed to provide MHS a guarantee that the files put into the archive are bit-for-bit the same over time. MHS will need a checksum of the files before moving them to the archive. The checksum will have to be kept locally, backed up and periodically validated against the files in the archive. This is required in Instrumental's opinion given that:

1. Though the cloud provider might state that file checksums are validated regularly, it is MHS that has the legal responsibility for the data stored with the cloud provider.
2. No software and/or hardware is perfect and MHS must validate the files placed into the cloud.

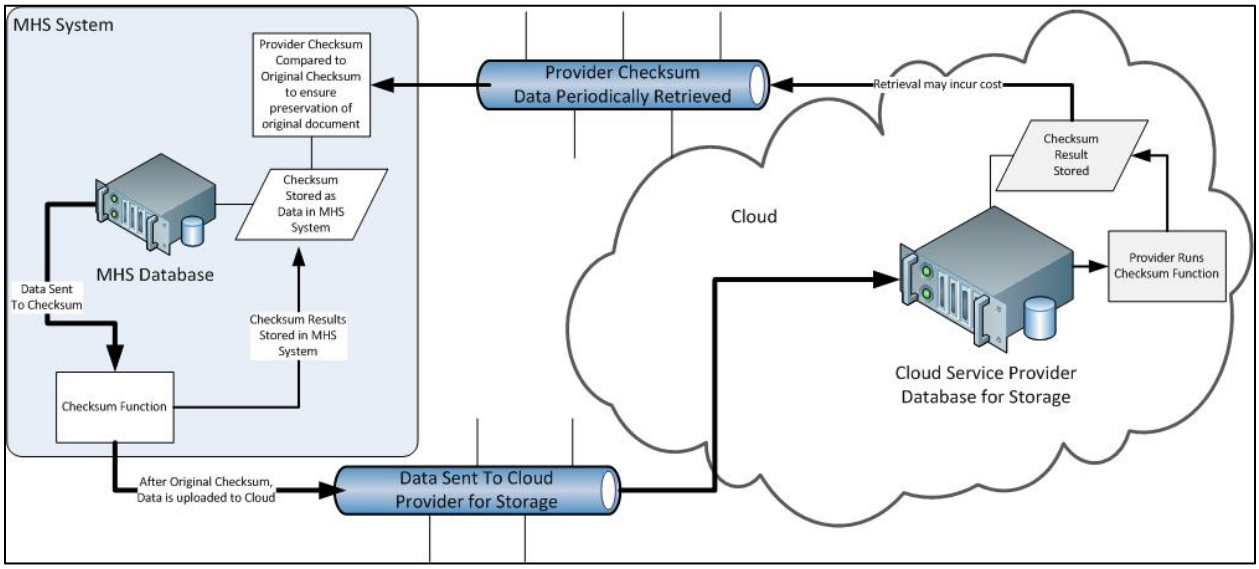

Figure 1 – Proposed Data Integrity Validation Procedure

Instrumental suggests that files should be randomly validated and that 1% of the files be validated monthly. It should be noted that any checksum failure might not be the fault of the cloud provider, but it could be the result of silent data corruption on the WAN. For this reason, the file should be read a number of times and TCP/IP error rates should be checked before failure is declared.

The following table summarizes the capabilities and cost information for the various vendors.

Provider / Service	Data Integrity	Reliability	Scalability	Retention and Portability	Availability	Data Ownership	Preservation Functionality	Total 3-Yr. Costs
Amazon S3	Limited checksums	Average	Almost unlimited	Not easy to move	Average	Similar to others	None	Medium
Amazon Glacier	Limited checksums	Multiple tape copies	Almost unlimited	Not easy to move	Lower since on tape	Similar to others	None	Low
Google Cloud Storage	No checksums	Average	Almost unlimited	Not easy to move	None in contract	Contract concerns	None	Medium
Tessella's Preservica	Checksums and CRC	Multiple cloud copies	Same as S3	Multiple providers	Multiple cloud copies	Similar to others	Developed for this	?
VISI ReliaCloud	Limited checksums	Average	Cannot support MHS	Somewhat limited	Average	Similar to others	Some claimed	High
SDSC Cloud Storage	Automatic verification	Average but no tapes	Almost unlimited	Claimed to be easy	Concerns about disks	Similar to others	None	Medium
DuraSpace DuraCloud	User run checksums	Multiple cloud copies	Same as S3	Multiple providers	Same as S3	Similar to others	Some claimed	Medium
IBM SmartCloud	Limited checksums	Average	Unknown	Unknown	None in contract	Similar to others	None	?
FujiFilm Permivault	Custom plans	On-site and cloud	Almost unlimited	Somewhat limited	On-site copy	Similar to others	Limited	Low
FujiFilm Permivault Client	Custom plans	Cloud only	Almost unlimited	Somewhat limited	Average	Similar to others	Limited	Low
Code 42 CrashPlan Pro	Limited checksums	Average	More limited than S3	Somewhat limited	Average	Similar to others	None	Low

Table 4 – Comparison of Capabilities and Pricing[15]

The legend below defines the colors used in the table above:

	Very good to best capability
	Average capability
	Poor or no capability
	Unknown data

[15] Additional information from Duracloud regarding pricing has led to a reclassification from the High cost category in the previous version of this report to Medium.

As can be seen from this table, some products do not provide some of the required capabilities (for example, the current lack of scalability of the VISI ReliaCloud product). Most of the cloud vendors provide no preservation functionality; if this is a requirement for the MHS collections, the cloud vendors without this capability should not be considered. There is a wide range of pricing as is shown in this table, but for the lowest cost options (CrashPlan Pro, Permivault client and Amazon Glacier), there will need to be some decisions on whether MHS can live with the somewhat limited capabilities of these products.

Appendix A – Acronym List

AES-256 - Advanced Encryption Standard 256 bit key size
API - Application Program Interface
AWS - Amazon Web Services
CPU - Central Processing Unit
DIP - Dissemination Information Package
DRA - Durable Reduced Availability
FTP - File Transfer Protocol
Gb/sec - Gigabits per second
GbE - Gigabit per second Ethernet
GiB/sec - Gibibyte (2^{30} bytes) per second
GUI - Graphical User Interface
HIPAA - Health Insurance Portability and Accountability Act
HPC - High Performance Computing
HTTPS - Hypertext Transfer Protocol Secure
HVAC - Heating, Ventilation and Air Conditioning
I/O - Input / Output
LDAP - Lightweight Directory Access Protocol
LOCKSS - Lots of Copies Keep Stuff Safe
LTO - Linear Tape-Open
LTFS - Linear Tape File System
MHS - Minnesota Historical Society
NAS - Network-Attached Storage
NDIIPP - National Digital Information Infrastructure and Preservation Program
NOAA - National Oceanic and Atmospheric Administration
OCR - Optical Character Recognition
PiB - Pebibyte (2^{50} bytes)
PCI DSS - Payment Card Industry Data Security Standard
REST - Representational State Transfer
SAN - Storage Area Network
SDB - Safety Deposit Box
SDSC - San Diego Supercomputing Center
SLA - Service Level Agreement
SOP - Standard Operating Procedure
SSAE 16 - Statement on Standards for Attestation Engagements, Number 16
SSD - Solid-State Drive
S3 - Amazon's Simple Storage Service
TiB - Tebibyte (2^{40} bytes)
TRAC - Transparent Approach to Costing
XML - Extensible Markup Language

Appendix B – Permivault Compatible Tape Drives and Software

STRONGBOX

Crossroads StrongBox® Interoperability Matrix

Release: StrongBox 1.7.5.35.0 Last Update: 12/1/12

Tape Libraries

Manufacturer	Library Model	Tape Drive	Comments/Qualifiers	Tested	Qualified
Dell	TL2000	HP LTO-5		✓	✓
HP	ESL G3	HP LTO-5	Library Firmware 602H.GS07601	✓	✓
HP	MSL 2024	HP LTO-5	Library Firmware 5.40	✓	✓
HP	MSL 4048	HP LTO-5	Library Firmware 6.90		✓
HP	MSL 8096	HP LTO-5	Library Firmware 100013.00E		✓
IBM	TS3100	IBM LTO-5	Library Firmware A.60	✓	✓
IBM	TS3200	IBM LTO-5	Library Firmware A.60	✓	✓
IBM	TS3500	IBM LTO-5	Library Firmware A.430	✓	✓
Overland	Neo4000	HP LTO-5	Library Firmware I334		✓
Qualstar	RSL-8350 (FC Only)	IBM-ULT3580		✓	✓
Quantum	Scalar i500	HP LTO-5	Library Firmware i7.3.1.621G	✓	✓
Quantum	Scalar i6000	HP LTO-5	Library Firmware 607A.G505301	✓	✓
Spectra Logic [1]	T50e	IBM LTO-5	Library Firmware BlueScale 12.0.3-20111122F	✓	✓
Spectra Logic [1]	T120	IBM LTO-5	Library Firmware BlueScale 12.0.11-20120509F	✓	✓
Spectra Logic [1]	T950	IBM LTO-5			✓
Spectra Logic [1]	T-Finity	IBM LTO-5	Library Firmware BlueScale 12.0.3-20111122F	✓	✓

Notes: The use of a FC switch is recommended. If a switch is not used, use L-Port (Loop Port) mode.
Notes: [1] If using Sprectra T-series library, disable Media Lifecycle Management (MLM).

Applications

Vendor	Application Name	Type	Tested	Qualified
Adobe	CS5 Premiere Pro	Non-linear Editing (NLE)	✓	✓
Adobe	CS6 Production	Non-linear Editing (NLE)	✓	✓
Apple	Final Cut Pro 7	Non-linear Editing (NLE)	✓	✓
CatDV	CatDV 9.0	Media Asset Manager (MAM)	✓	✓
Empress	eMAM	Media Asset Manager (MAM)	✓	✓
Caminosoft	Managed Server HSM	Archiving/HSM	✓	✓
Crossroads Systems	FileStor HSM	Archiving/HSM	✓	✓
EMC	DiskXtender (6.4), (6.5)[1]	Archiving/HSM	✓	✓
Hitachi Data Systems (HDS)	NAS Filer	Archiving/HSM	✓	✓
Moonwalk	Moonwalk	Archiving/HSM	✓	✓
Symantec	Vault 9	Archiving/HSM	✓	✓
Apple	Finder	File / Data Management	✓	✓
NTP Software	QFS	File / Data Management	✓	✓
LikeMAC	DiskOrder	File / Data Management	✓	✓
StoredIQ	DataIQ	File / Data Management	✓	✓
Windows	Explorer	File / Data Management	✓	✓
Dell (Qwest Software)	vRanger 6.0 [2]	Virtual Machine Backup	✓	✓
NTP Software	Precision Tiering (ODDM component)	Storage Tiering / Data Migration	✓	✓
PoINT	Storage Manager	Storage Tiering / Data Migration	✓	✓
Docuprotection	Docuprotection	Secure File Storage	✓	✓

Notes: [1] 6.5 support is limited to files w/o alternate data stream (fix in StrongBox 1.7.5 release).
Notes: [2] Solution requires the use of the StrongBox client (FileStor-HSM) to manage data movement to the StrongBox.

Host OS Connectivity (Mountable Shares)

Vendor	Application Name	Protocol	Comments/Qualifiers	Tested	Qualified
Apple	MAC OS X	CIFS, NFS	Snow Leopard (10.6), Lion (10.7), Mountian Lion (10.8)	✓	✓
Open Source	Linux CentOS	CIFS, NFS	Versions 5.5, 6.0, 6.1	✓	✓
Microsoft	Windows 7	CIFS	Version Ent 64-bit	✓	✓
Microsoft	Windows XP, XP Pro	CIFS	Version SP3 32-bit	✓	✓
Microsoft	Windows Server 2008	CIFS	Version R2 64-bit	✓	✓

Notes:

Web Browsers
The following browsers are supported for use viewing the StrongBox web interface:
Note: Javascript and cookies must be enabled.

Vendor	Application Name	Comments/Qualifiers	Tested	Qualified
Apple	Safari	Version 5.1.7, 6.0; Mac OSX 10.6 and 10.7, 10.8	✓	✓
Google	Chrome	Version 14.0.835		✓
Google	Chrome	Version 23.0	✓	✓
Microsoft	Internet Explorer	Versions 7.0		✓
Microsoft	Internet Explorer	Versions 8.0, 9.0	✓	✓
Mozilla	Firefox	Version 13.0.1 and 14.0.1	✓	✓