# Recommendations Regarding a Third HathiTrust Instance

## Introduction

HathiTrust (www.hathitrust.org) is a shared preservation and access repository managed by a consortium of major US academic research libraries. The design and operation of the repository are predicated on fulfilling HathiTrust's mission "to contribute to the common good by collecting, organizing, preserving, communicating, and sharing the record of human knowledge."

Within the preservation community there is now widespread acceptance that *redundancy* – and in particular, highly *de-correlated* redundancy – is one of the most significant technical strategies for ensuring the long-term usability of managed digital content, primarily by guarding against disruptive, and potentially destructive, single points of failure.

Thus, desirable redundancy is best provided through a combination of:

- Fault-tolerant hardware with redundant data paths, controllers, power supplies, etc.
- Geographical replication.
- Technological heterogeneity.[1]

HathiTrust currently operates in a mirrored, two-instance configuration, based on Isilon storage hardware and replication software (www.isilon.com).[2] The storage at each instance provides localized redundancy in the form of fault tolerant hardware (dual data paths, controllers, power supplies, etc.) and the use of *n*+3 RAID, which protects against a failure of up to three disks in each RAID group.[3] Nevertheless, this redundancy is only partially de-correlated since the storage, and its supporting infrastructure (power, cooling, network connectivity), is collocated within a data center. Global de-correlation is provided by having all managed content mirrored at two independent instances. While this two-node redundancy is clearly beneficial, it is open to (constructive) criticism along the following lines:

- Two nodes do not provide a means to "break the tie" if the replicas are ever found to disagree with one another.
- Periodic site-wide maintenance or interruption of service at one of the nodes results in the total loss, however brief, of the benefits of geographic redundancy.
- Technology homogeneity, which leaves both nodes vulnerable to systemic problems specific

[1] It is important to note that although technological heterogeneity increases overall preservation assurance, it also raises ongoing operational costs due to factors such as the necessities for parallel training and procedures, and fewer opportunities for exploiting administrative economies of scale.

[2] More accurately, HathiTrust maintains two *on-line* replicas (on-site magnetic disk) along with a third *off-line* copy (off-site encrypted magnetic tape). However, the tape replica provides a lower level of function and preservation assurance than the disk copies, primarily with regard to access responsivity (e.g. on-line vs. off-line), transparency (encryption), external dependencies (encryption key management), and control (partner-operated data centers vs. commercial data center). The external digitization vendors (Google and Internet Archive) are a potential source of a fourth replica, although the circumstances under which this data could be re-acquired by HathiTrust are uncertain and neither vendor will hold more than a subset of the complete HathiTrust holdings.

[3] There is no standard nomenclature within the storage community for *n*+3 RAID (three parity disks per RAID group), although it is the natural technological successor to RAID 5 (*n*+1) and RAID 6 (*n*+2). The use of *n*+3 RAID is an Isilon-recommended best practice.

to the Isilon systems and software.

In light of these factors, the HathiTrust Executive Committee convened an ad-hoc working group (see Appendix A) to investigate the question of a creating a third repository instance.

## Analysis

Since more instances of the repository provide higher levels of preservation assurance than fewer, a third instance is obviously advantageous and desirable *all things being equal*.[4] But of course, things are never equal and the self-evident benefits of a third instance must be weighed carefully against its concomitant costs, including:

- The capital cost of hardware/software acquisition.
- The ongoing cost of vendor support.
- The cost of hardware hosting, including data center rack space, power, cooling, network connectivity, monitoring, etc. (This cost could be obviated somewhat by exploring cloud hosting options.)
- The cost of additional staff resources. (Again, this could be minimal with a cloud-based solution.)
- The cost of technical heterogeneity, including duplicative training, system administration, replication workflows, ongoing capacity provisioning, media refresh, etc.
- The cost of additional organizational complexity.
- The (somewhat intangible) cost of added conceptual complexity of the overall replication architecture.

The balancing of the cost and benefit of a third instance is also complicated by the fact that, as is often the case in technical decision making, most of the costs are immediate and tangible while the benefits are somewhat intangible and accrue slowly over time.

The charge from the Executive Committee lays out a number of important technical, managerial, and organizational questions, including:

- Is platform diversity necessary or desirable?

  Again, all things being equal, platform diversity is obviously desirable. A technological monoculture is susceptible to systemic failure that could, in the extreme, lead to unrecoverable data loss, as in the case of simultaneous system-specific failure at all replicas. Platform diversity significantly reduces the likelihood of such an event.

  The HathiTrust partner institutions possess a broad gamut of technological expertise, so HathiTrust is particularly well-suited to accommodate heterogeneity in terms of managing multi-vendor, geographically dispersed systems. However, due to a lack of vendor-neutral standards for data synchronization, the management of synchronization between technologically heterogeneous systems, particularly at the petabyte scale at which HathiTrust operates, will require customized solutions that are complex and costly.
- Should redundant content be open, portable and interoperable?

[4] Although the "more is better" strategy is sound in principle, there is undoubtedly a point of practical diminishing return, but only at a much greater degree of replication.

The answer to this question seems particularly self-evident: it is hard to imagine under what circumstances HathiTrust would want or accept having content in a form that was closed, location-specific, or single-purpose. Long-term preservation clearly requires the ongoing evolution and replacement of technology and strategy, activities that are greatly facilitated by open, portable, and interoperable solutions.

- How important is geographic distribution? What constitutes sufficient geographic distribution?

  Geographic distribution increases the de-correlation of the replicas, decreasing the likelihood that a problem at one replica would affect the others. This is particularly true for natural disasters (fire, flood, tornado, earthquake, etc.) and infrastructure failures (power, network, etc.).

- What capabilities does a vendor need to exhibit to be considered as a viable candidate for a third instance?

  Other than stating that a non-Isilon platform is welcome for purposes of technological heterogeneity, the desirable characteristics of potential vendors are probably the same as those for any other significant technical acquisition, including:

  - Sound and transparent technical solutions;
  - Good technological and support track record;
  - Demonstrated organizational stability and evidence of longevity; and
  - Competitive pricing or, even better, a willingness to provide significant no-cost equipment grants.

- What is the appropriate medium for a third copy? Is cloud storage a good candidate for a redundant copy?

  The question of media is highly dependent on the expected use cases for the replica. If the replica is intended as a bright archive providing load balanced access to managed content, then magnetic disk storage is necessary. On the other hand, a dark replica provided solely for additional preservation assurance could be implemented at lower cost on magnetic tape.

  Other factors to consider in the disk to tape comparison include:

  - While tape storage is currently less expensive then disk, the per-unit price of disk seems to be dropping at a greater rate than that for tape, suggesting that at some future point the price advantage will change over.
  - Due to its inherent random access modality, disk storage lends itself better to message digest-based auditing than tape. Periodic auditing to insure the bit-level integrity of managed content is an important component of preservation strategy.
  - Media refresh for tape is more time consuming than that for disk.

  While non-magnetic storage choices are available, such as optical disk and film (digital and analog), their use by the preservation committee is limited and a clear understanding of the positive and negative implications regarding their long-term use is fragmentary

  Cloud storage provides a number of advantages by avoiding costs associated with local

physical hosting (data center costs, training, support, etc.). However, the use of a cloud hosting arrangement would have to be accompanied by a service level agreement under which the service provider accepts a meaningful obligation to maintain and make available the stored content in perpetuity without loss, and provide adequate security measures as required by HathiTrust's own contractual obligations. Anecdotal evidence suggests that no providers currently offer such service level agreements.

In addition to the questions explicitly posed by the Executive Committee, some other pertinent considerations include:

- Is it necessary to consider the third instance as a permanent replica?

  As was mentioned previously, one concern arising from the current two node architecture is the loss of redundancy during periods of node-wide service maintenance or other interruptions to normal service. HathiTrust may want to consider the establishment of transient replicas during times of increased preservation risk such as media refresh or platform migration, or during periods of recovering from unplanned but sustained service outages. The use of cloud-based storage solutions would be the most effective and affordable option for this temporary storage need.

- What effect will a third instance have on the existing HathiTrust partnership agreement and business model?

  Although an important consideration, this is a question that lies outside the area of expertise of the working group, and is better considered by the Executive Committee or some other appropriately constituted group.

One additional point that is important to raise is the economic factor. The current analysis is somewhat incomplete in the sense that no assumption can be made about the economic basis under which a third instance could be acquired. Even at normal academic pricing levels, the necessary hardware and software acquisition would probably be considered cost prohibitive. While a no-cost grant of equipment would obviously be welcome, ongoing support, hosting, and administrative costs would still be non-trivial.

The attached spreadsheet (ht-3rd-instance-decision-factors.xls) provides a quantitative summary of the working group's analysis. Decision factors were developed in three categories: preservation, access, and operational cost. Raw evaluation scores are normalized by weights reflective of the relative importance of the various factors. Seven separate scenarios are considered:

- Hosted instance based on Isilon disk storage appliance
- Hosted instance based on non-Isilon disk appliance
- Hosted instance based on commodity disk
- Cloud instance
- Hosted instance based on virtual tape library
- Hosted instance based on conventional tape library
- Not establishing a third instance

There is a clear preference for establishing a third instance, but it is important to note this evaluation model may underemphasize the importance of economic considerations. Assuming a third instance, there is a preference for a locally-hosted disk-based instance over cloud or tape, with a slight preference for provisioning the instance with non-Isilon equipment.

Still, the pertinent question remains: should HathiTrust establish a third instance?

## Summary

First, it must be stated clearly that the reliability of the existing HathiTrust architecture is very high, so a third repository instance is *not necessary* in order for HathiTrust to meet its preservation obligations. Nevertheless, a third instance would provide a tangible *increase* to HathiTrust's overall level of preservation assurance, therefore it remains a desirable goal. That desire, however, must be tempered by economic reality and the Executive Committee must grapple with the issue of how much assurance they are willing to buy, or conversely, how much risk they are prepared to accept.

Assuming the acquisition of a third instance at normal academic pricing, the working group believes, based on their collective experience and intuition, that the costs of acquisition and operation would be proportionally greater than the marginal increase in preservation assurance. This should not be interpreted as deprecating the preservation value of a third instance, but merely that the cost/benefit equation is not favorable at this time.

## Recommendations

1. Given the high level of preservation confidence in the existing two instance architecture, and absent specific favorable economic terms for acquisition and operation, there is no urgency in establishing a third repository instance at this time.
2. However, a third instance would provide small, but nevertheless tangible preservation benefits and should be considered if it could be established and operated on favorable economic terms.
3. Since the level of external support for establishing a third instance is unknown (for example, vendor discount or grant funded support), HathiTrust should be prepared to respond quickly to opportunities to establish a third instance on favorable economic terms as they may arise in the future.

## Appendix: Ad-hoc Working Group

The ad-hoc working group created by the HathiTrust Executive Committee to investigate and make recommendations regarding a third instance of the HathiTrust repository was composed of the following members:

- Stephen Abrams — California Digital Library (co-chair)
- Luc Declerck — University of California, San Diego
- John Kunze — California Digital Library (co-chair)
- Robert Lowden — Indiana University
- David Minor — San Diego Supercomputing Center
- Cory Snavely — University of Michigan
- John Towns — NCSA/University of Illinois

The group conducted its investigation during August through November, 2009.