

Archives Storage Solutions Technical Final Proposal

Prepared by: Randy Spainhower and Joe Tripp

This is a final proposal for Archives Storage needs in FY 2016 and on. Each Storage Space outlined in the Archives Data Storage Business Analysis (Prepared by Gerry Satterlee) is discussed in the subsequent sections:

Quarantine Store

As Archives acquires records they first store them in a Quarantine space for 30 days. Right now they are using external hard drives for this purpose. It is our recommendation that they continue with this practice as this area really can be “quarantined” by storing the external hard drive unconnected for 30 days. Plus they have a second copy of the data from the initial submission.

Local Processing Store

Archives has need for a place to put submission data after it has been quarantined for 30 days. This Local Processing Store will be an area that has low network latency and can handle the workload of using Photoshop to convert digital master images (usually uncompressed TIFFs) to a format that's easier for the public to download, like a JPEG. Sometimes hundreds of thousands of images in a matter of days or weeks need to be converted. At the same time they might crop it or improve the contrast, all in the batch action that's available in Photoshop. In light of the work that they do with this part of the Process DTS is recommending placing a NAS(Network Attached Storage) device local at the Division of Archives office. This device will be able to accommodate the needs of the Local Processing Store and the Transitory Digital Image Store(See corresponding section). Archives will pay the published NAS rate to DTS for allocated storage, which currently is \$0.10/GB/Month. Because this is a temporary space during processing the Division of Archives requested no backup or replication for Local Processing Store. See below for costs of 10 TB allocated according to the Business Analysis Scale section:

Proposed Cost

NAS storage cost is \$0.10/GB Allocated

10 TB Share = \$1,024/Month, \$12,288/Year

Preservation Data Store

The Archives Preservation system is the final store for permanent archival digital data. The data needs to maintain it's checksum (fixity) for the long haul. Presently the Preservation System includes a server hosting the archiving application (AXAEM) and tools, the HNAS low cost storage houses the preservation data and is co-located (replicates to RF). It is our recommendation that at present and for the near future the existing Enterprise HNAS will meet

the needs of the Preservation System for the main and secondary copies of data. Division of Archives has expressed that they need a 3rd copy on a different media type for their data resiliency requirements. DTS proposes placing an HP Tape Library at the Division of Archives office attached to the NAS device with LTFS drivers. This can be used to make a 3rd copy of the Preservation data, tape cataloging and library management will be handled by Division of Archives staff in line with business processes. Division of Archives will be responsible for costs associated with procurement, support, and maintenance of the Tape Library. If DTS is needed to consult on the HP Tape Library, LTFS interface, or Tape the Division of Archives will pay on a time and materials basis. The current DTS Consulting Hourly Charge is \$77.87/Hour. Division of Archives will be responsible for all purchase of all Tape Media for this solution as well. Current costs of Preservation Data are below. This is just for storage costs of preservation data and don't include the application server and web server costs which are minimal as most data is on the HNAS. The Proposed Tape Library budgetary costs are below as well.

Current costs

HNAS storage cost is \$0.10/GB Allocated

206 GB Share = \$20.56/Month, \$246.72/Year

206 GB Replicated to Richfield as backup = \$20.56/Month, \$246.72/Year

Total Cost = \$41.12/Month, \$493.44/Year

Allocation can be increased at an interval of 250GB. Each GB of allocation is \$0.10 a month which equals each TB allocated being roughly \$102 a month.

Proposed Tape Library with LTFS technology

HP MSL6480 Tape Library with 2 LTO-6 Tape drives and 80 LTO-6 Tape Media, 5 years of 24/7 support and maintenance included

Budgetary Quote = \$48,274.64 (Division of Archives will reimburse DTS for this expense)

Access Copy Store

Archives has multiple places they might put access copies, these include: CONTENTdm (hosted at OCLC), local webservers (archives.utah.gov and openrecords.utah.gov), as well as locally stored M-Discs indexed for reference. DTS is not recommending changes to the CONTENTdm or local web server storage. However the Division of Archives has expressed that they may have need for an Optical Jukebox for M-Disc Blu-Rays to help automate and increase productivity of this process in the near future. At this time it seems that this is outside of this initial push for cost reasons. However DTS has identified a few candidates for such a solution when Archives is ready for this process to be implemented. DTS would purchase the Optical Jukebox and place it at the Division of Archives office in a staff assigned location. Archives will then reimburse DTS the amount of the purchase. Archives will pay the network services port charge for the device of \$44/Month and the Security rate (\$15/Month) for the device. Archives will also purchase any media(M-Discs) needed. If archives staff need support for the device

from DTS after installation, it will be under the DTS Consulting charge of \$77.87/hour plus materials.

Migration Space Store

Migration space is on a case by case basis when Archives needs to do a change of some data in the Preservation System already. DTS has SAN and HNAS storage that can be allocated temporarily for such an operation. It is our recommendation that nothing be dedicated to this space until a migration operation is needed. This will ensure that Archives is not paying for unnecessary storage space. When needed they can request the amount of space they will need and we can allocate the migration space for the duration of the migration and Archives will only pay for the time they need the space for. They can request high speed SAN or low cost HNAS as their respective rates.

Transitory Digital Image Store

This space is used as an external working space to store scanned images until staff can do the final work on them. The Scanners and corresponding computers are kept offline in a peer-to-peer network as recommended by the vendor.

In light of the work that they do with this Process, DTS is recommending placing a NAS(Network Attached Storage) device local at the Division of Archives office. This device will be able to accommodate the needs of the the Transitory Digital Image Store and the Local Processing Store (See corresponding section). This will be accomplished by having separate physical network links to the device so that the two separate Data Stores can stay separate on their existing networks. A dedicated link will go to each area but the data for both will be maintained in one device.

Archives will pay the published NAS rate to DTS for allocated storage, which is \$0.10/GB/Month. It had been suggested by Archives staff that a secondary copy could be put on the local LTFS Tape Library as the backup, but because it is a manual process to put a tape in and copy it DTS strongly discourages that approach. DTS recommends that a copy will be replicated as a backup to one of DTS's main Data Centers and the same rate applies for the replicated copy of \$0.10/GB/Month. But the decision on which process to use for backing up the Transitory Digital Image Store will be decided by the Division of Archives. See below for costs of 10 TB allocated with replication according to the Business Analysis Scale section:

Proposed Cost

NAS storage cost is \$0.10/GB Allocated

10 TB Share = \$1,024/Month, \$12,288/Year

10 TB Replicated to SLC or Richfield as offsite copy/backup = \$1,024/Month, \$12,288/Year

Total Cost = \$2,048/Month, \$24,576/Year

BY SIGNING BELOW, YOU CERTIFY THAT YOU HAVE READ THIS PROPOSAL/AGREEMENT, THAT YOU ARE RESPONSIBLE FOR COSTS INCURRED WITH THIS PROPOSAL, THAT YOU KNOW AND UNDERSTAND THE MEANING AND INTENT OF THIS AGREEMENT AND THAT YOU ARE ENTERING THIS AGREEMENT KNOWINGLY AND VOLUNTARILY.

By: _____
Patricia Smith-Mansfield

Title: Director Division of Archives

Date: _____

Appendix

Data Transfer Requirements

Because there will be multiple processes utilizing the Wide Area Network (WAN) connection for ingest to the Preservation System, ContentDM, and backup of Transitory Digital Storage, there is a need to determine whether the existing WAN circuit can support the data transfer requirements. It has been confirmed that the current WAN circuit will support up to 200 Mbps. With circuit latency and contention with other network traffic, it is estimated that the circuit should be able to consistently support approximately 100 Mbps of data transfer. Using the following formula, we estimate that the circuit can support a data transfer of the following calculated value during the daily 6:00pm to 6:00am time period:

Formula: *Average throughput in Gigabytes per second x Number of seconds in time period*

$100 \text{ Megabits} / 8 = 12.5 \text{ MegaBytes} * 1000 = .0125 \text{ GigaBytes}$

$12 \text{ hours} = 720 \text{ minutes} = 43200 \text{ seconds}$

Estimated Nightly Data Transfer Capacity = .0125 GB x 43200 seconds = 540 GB / 12 hours

The WAN circuit should be able to support the transfer of approximately 2.5 TB during a 60 hour weekend period.