The Bit Mountain Research Project

By Shane Hathaway and the Touchstone Team

January 2006

Table of Contents

1 Project Goal

The Church intends to build a large repository of digital images of value to family history research. The goal of the Bit Mountain research project is to discover and evaluate the risks the Church faces in building a reliable 18 petabyte digital repository, and learn how to manage those risks. We performed this research by studying media options and building software for managing a MAID (massive array of idle disks) repository. The experience of building and testing the software has taught us a lot of what we needed to learn. It has also produced software that may help solve the problem.

1.1 Requirements

1. **Store 18 petabytes.** The Granite Mountain Records Vault is working to scan all of its microfilm into digital format. The FamilySearch Scanning project has already built a system that scans microfilm quickly with minimal human intervention, but that project does solve the problem of retaining the scanned images. The vault estimates they have 3 billion images to scan and that the average size of a preservation-quality image is 6 MB, totaling 18 PB.

   Also, the Digital Processing Center is building an inventory of digital-born images to preserve. By the end of the 2006, the DPC expects to produce 1.5 TB per day. Eventually, the DPC is likely to have larger storage requirements than the vault has.

2. **Store Reliably.** We need to preserve all images at least until the next major upgrade in storage technology. Current high density storage technology has an average shelf life of 1 to 5 years. Storage technology designed for archival has an average life of 30 years. While industry is moving toward higher density but less reliable storage, our images need to survive through the millennium. Since we don't know how much time will pass before high density storage technology becomes more reliable, we need a high assurance that the data we store now will outlive the media that stores it.

3. **Minimize system administration.** Storage vendors estimate that typical petabyte-sized storage farms require a storage administrator for every 200-300 TB. Administrators spend their time replacing hard drives and reconfiguring the system to use the new drives. If we hired the 50 to 90 full time administrators required to handle 18 PB, their salaries and benefits could dwarf all other costs over time. An ideal system would only require occasional attention from a team of two or three system administrators who would simply replace failed media.

4. **Minimize hardware, software, power, and infrastructure costs.** Commodity hard drives sell for around 50 cents per gigabyte (which comes to $9 million for an 18 PB repository) as of this writing, yet typical disk storage vendors want $5-10 per GB (which comes to $90-180 million for an 18 PB repository.) The extra hardware and software they sell is primarily designed to maximize I/O transfer rates and availability, but this project does not have those requirements. We can tolerate downtime as long as the images become accessible later.

   Also, large storage arrays have major electrical power requirements, leading to major cooling requirements. We don't want a 5 or 6 digit power bill every month.

5. **Distinguish distribution requirements from preservation requirements.** The system we intend to build will have two digital repositories: a preservation repository, 18 PB in size, with lossless images on slow but highly reliable media; and a distribution repository, one tenth the size of the preservation repository, with lossy images on fast but less reliable media. The distribution repository will provide images for distribution to genealogists. Images lost from the distribution repository will be restored using the preservation repository. This design allows us to drop Internet-level speed requirements from the preservation repository.