Using Rosetta, StorageGRID, and New IBM Tape Solutions to Implement State-of-the-Art Digital Preservation

Gary T. Wright

wrightgt@ldschurch.org

ELUNA 2012

May 2012

image courtesy of IBM

Introducing the Church

- The Church of Jesus Christ of Latter-day Saints
- Global Christian church with 14 million members
- 700,000 students enrolled in religious training
- 3 universities, 1 college
- State-of-the-art audio-visual capabilities
- Scriptural mandate to keep and preserve records

photo by Henok Montoya

Church History Department

- Preserves records of enduring value from Church leaders, departments, universities, and affiliations (more than 35 organizations)
- Helps each organization develop a records management plan
- Typically, less than 10% of records are archived

Church History Library on Temple Square

Church Audiovisual Capabilities

- Audiovisual records will consume vast majority of archive capacity

Mormon Tabernacle Choir and Orchestra

Free Bible videos from biblevideos.lds.org

- 100+ PB in a decade for a single copy!

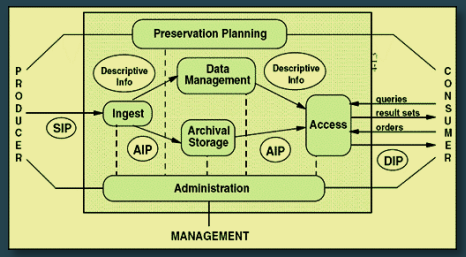
Conference Center on Temple Square

Architecting a Preservation System

- National Library of New Zealand business requirements

NATIONAL DIGITAL HERITAGE ARCHIVE PROGRAMME

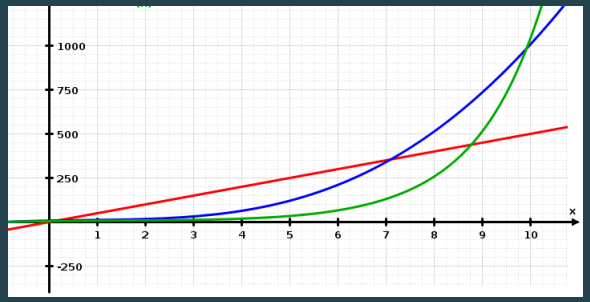
- OAIS Reference Model



- Minimize cost of archival storage

image courtesy of NetApp

- Scalability



Minimizing Cost of Archival Storage

Total cost of storage ownership study

- Over ten years, ownership and operating costs of tape are one third of associated costs for disk arrays
- Power and cost per TB advantages of tape are expected to increase over time
- Conclusion—for now, tape is required to sustain a multi-PB digital archive
- But . . . tape presents some challenges

Scalability to Meet Church Needs

- After research, scalability concerns eliminated open source repositories
- Ex Libris Rosetta SPOC
- White paper at exlibrisgroup.com
- Conclusion—Rosetta can meet Church needs
- CHIPS—Church History Interim Preservation System

DRPS Architecture

Digital Records Preservation System

Fixity Creation	DRPS Ingest Tools
Preservation Functions	ExLibris Rosetta
Fixity Bridge	Storage Extensions
Information Lifecycle Management	NetApp StorageGRID
Tape Interface	IBM Tivoli Storage Manager

Ensuring Archive Data Integrity

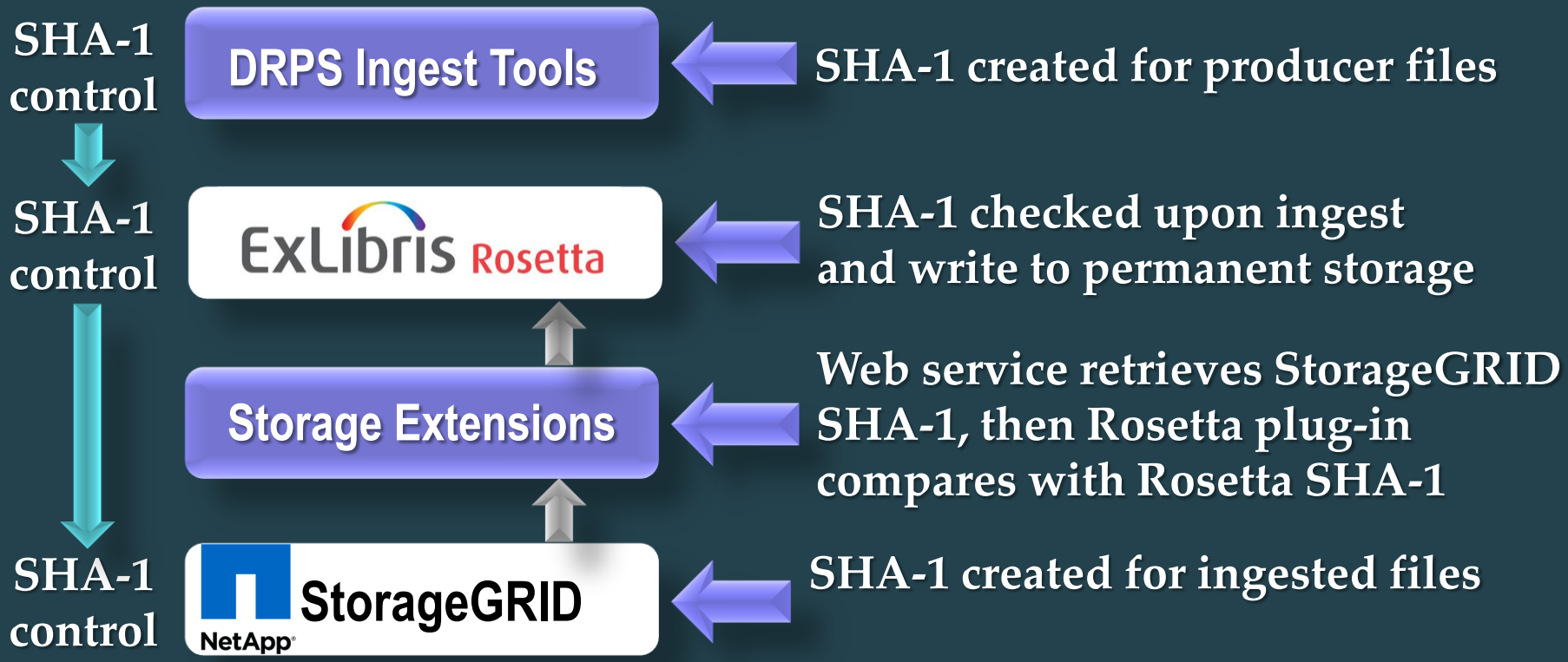
- A critical requirement of digital preservation—*differentiates an archive from other tape farms*
- Fixity information (secure hash checksum) enables data integrity validation by checking it every time data is written, transferred, moved, or copied
- End-to-end fixity checking should be performed from file ingest to permanent storage to delivery
- Periodic validation of the *entire* archive should also be performed to detect bit flips (i.e., bit rot)
- DRPS uses a variety of integrity values for fixity

Sizing Bit Errors

- Recent validation of entire DRPS archive resulted in a 3.3×10^{-14} bit error rate
- USC Shoah Foundation Institute visit

- 8 PB tape archive of videotaped interviews of Holocaust survivors and other witnesses
- Experienced 1500 bit flips in 8 PB (2.3×10^{-14} bit error rate)
- Real life measurements that provide guidance for tape archives

DRPS Data Integrity Validation



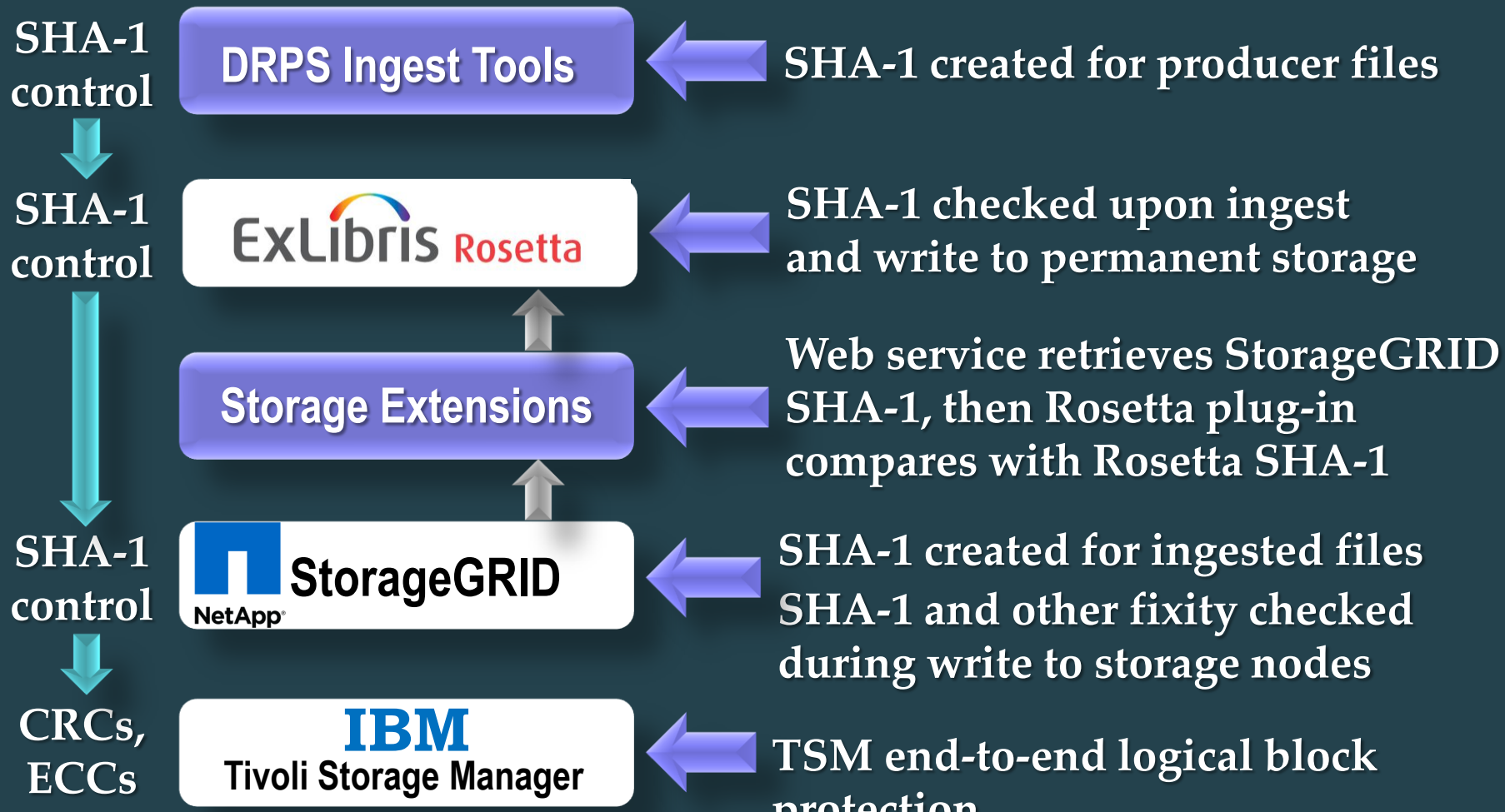
StorageGRID Fixity Checking

- StorageGRID is constructed around the concept of object storage

- Provides a layered/overlapping set of protection domains to guard against object data corruption

1. SHA-1 object hash—checked on store and access
2. Content hash—checked on access
3. CRC checksum—checked with every operation
4. Key-based hash value—checked on access

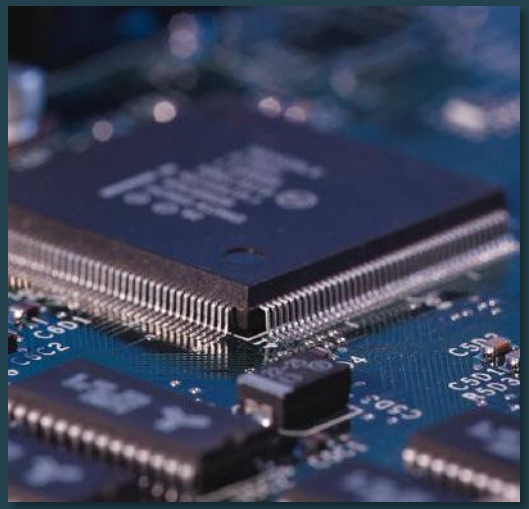
DRPS Data Integrity Validation



TSM End-to-End Logical Block Protection

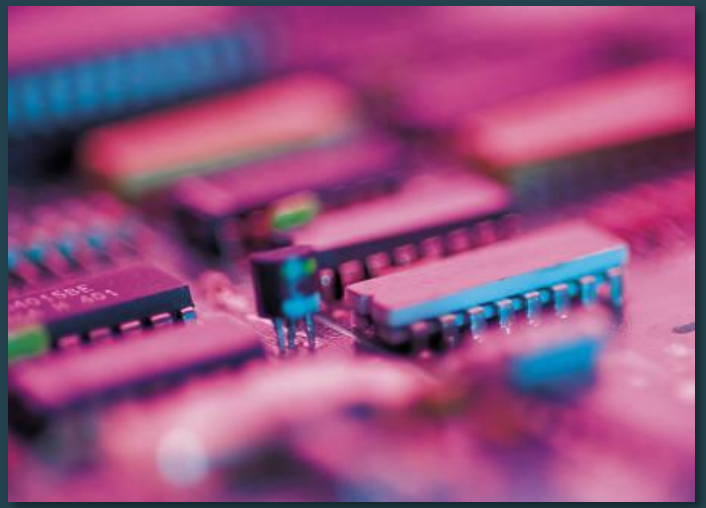
- Supersedes SHA-1 fixity information with cyclic redundancy check values (CRCs) and error-correcting codes (ECCs)
- Enabled with new, state-of-the-art functionality of IBM LTO-5 and TS1140 tape drives
- Seamlessly extends validation of data integrity until AIPs are correctly written to tape

TSM End-to-End Logical Block Protection

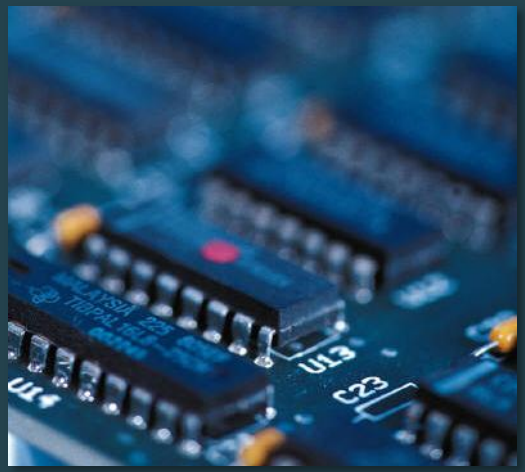


1. TSM server calculates and appends “original data CRC” to AIP logical block

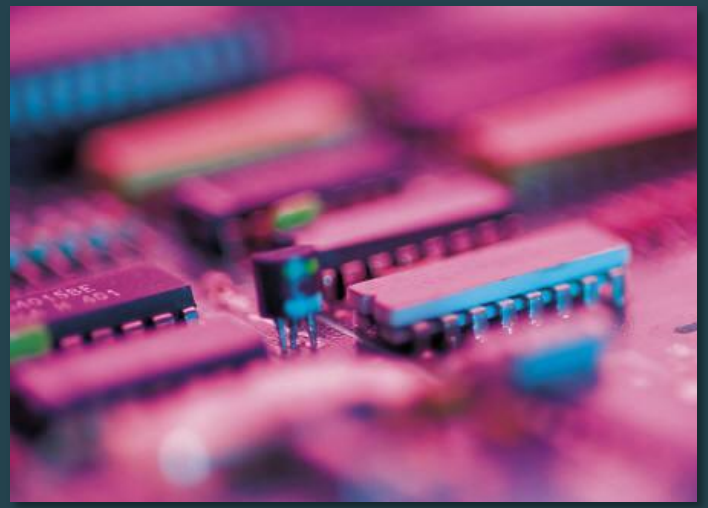
2. Tape drive computes its own CRC and compares to original data CRC



TSM End-to-End Logical Block Protection

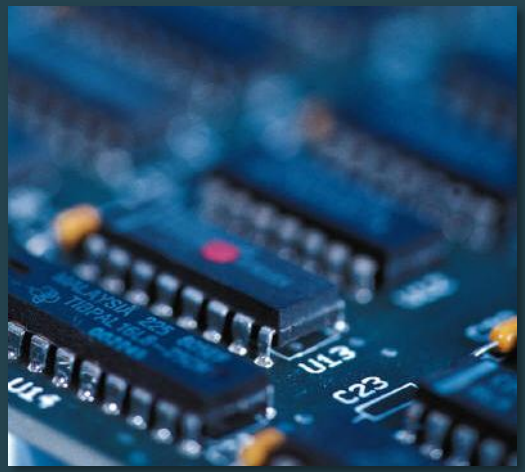


3. As logical block is loaded into drive data buffer, on-the-fly verifier checks original data CRC

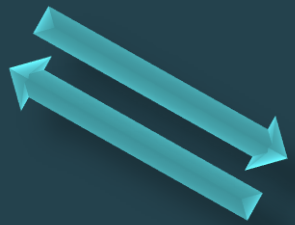


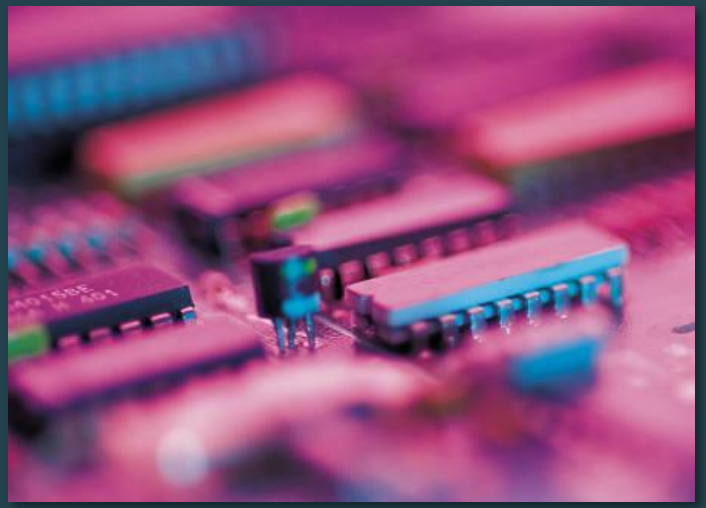
4. In parallel, a “C1 code” (Reed-Solomon ECC) is computed and appended

TSM End-to-End Logical Block Protection



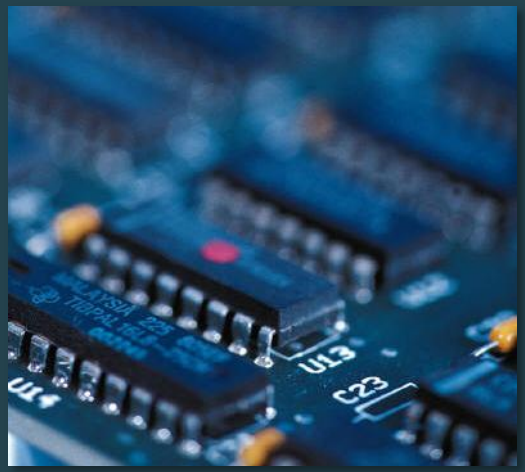
5. An additional ECC, referred to as "C2 code," is added to the logical block



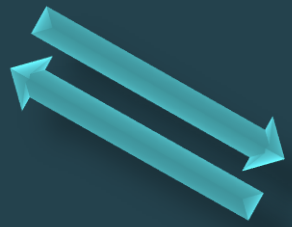


6. More powerful than the original data CRC, the C1 code is checked every time data is read from the buffer

TSM End-to-End Logical Block Protection

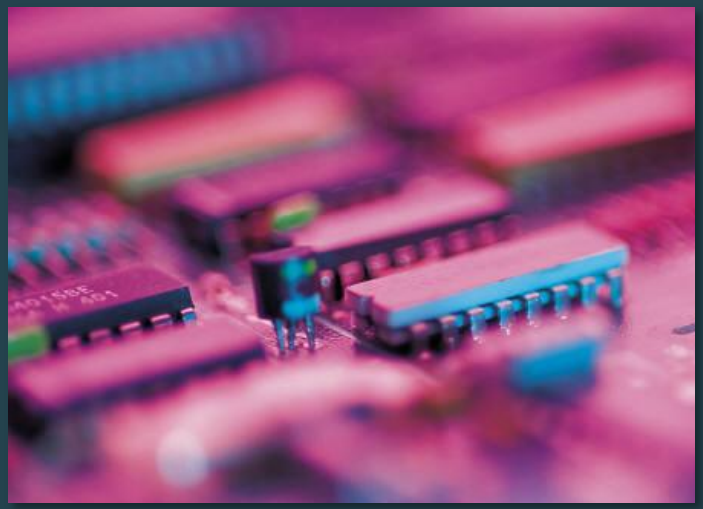


7. Data written to tape at full line speed with read-while-write process
8. Just written data loaded to buffer and C1 code checked



Successful read-while-write operation assures no data corruption from TSM client to tape

TSM End-to-End Logical Block Protection



9. When tape is read, all codes (C1, C2, original data CRC) are checked by drive

10. Original data CRC appended to logical block

11. TSM server verifies origin data CRC, completing TSM end-to-end logical block protection cycle

Ongoing Archive Data Integrity

- Must assume that bits will flip after being written correctly to tape

1001101101000001011

- Actual validation results indicate a 10^{-14} bit error rate

- Therefore, *all* tapes must be read periodically to identify and correct bit errors

image courtesy of IBM

Ongoing Archive Data Integrity

- Staging AIPs so a server can verify integrity is resource intensive!
- IBM LTO-5 and TS1140 tape drives provide a more efficient solution
- During "Verify" operation, a tape is mounted, drive checks all codes (C1, C2, original data CRC) as data is being read (at full line speed)
- Only status is reported as these internal checks are completed

image courtesy of IBM

DRPS Data Integrity Summary

- Fixity information is the key to data integrity
- SHA-1 values ensure data integrity to StorageGRID
- TSM end-to-end logical block protection ensures data integrity to tape
- In-drive validation enables ongoing integrity checks for the entire archive

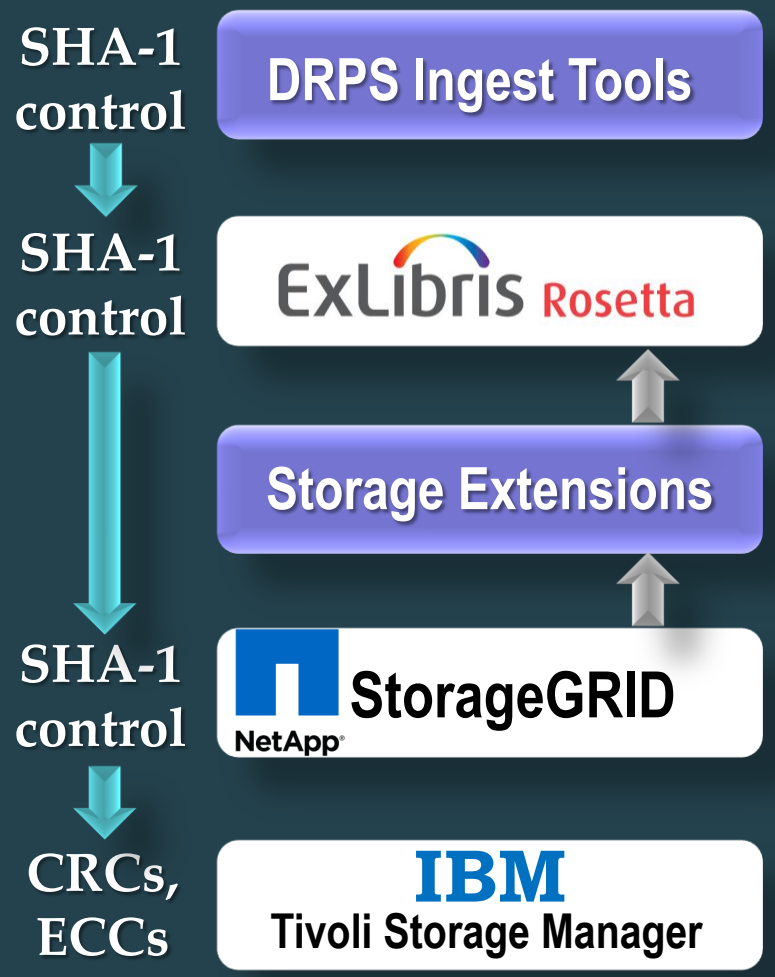


Image Preservation File Formats

- Lossless JPEG 2000 for still images (50% - 60% archive capacity savings with no loss of resolution)
- Motion JPEG 2000 wrapped in MXF containers for HD video files

Parish or Township of St. Thomas — Ecclesiastical District — City or Borough of New Salum — Town of — Village of

No. of House inhabited	Name of Street, Place, or Road, and Name or No. of House	Name and Surname of each Person who abode in the house, on the Night of the 30th March, 1851	Relation to Head of Family	Condition	Age of Males	Age of Females	Rank, Profession, or Occupation	Where Born	Whether Blind or Deaf and Dumb
72	High Street	Eliza Morris	Daug	U		17	Dressmaker	Wilts Salesbury	
	nowan yard	Sarah Anne Dº	Daug	U		16		Dº Dº	
		Emma Elizabeth Dº	Daug			13	Scholar	Dº Dº	
73	High Street	George Jordex	Head	Mar	62		Boot and Shoe maker	Dº Dº	
		Mary Ann Dº	Wife			58	Dº & Dº Dº Assistant	Dº Dº	
		Sarah Dº	Daug	U		17	Apprentice to Milliner	Dº Dº	
		Emily Dº	Daug			12	Netting	Dº Dº	
74	High Street	Benjamin Blake	Head	Mar	36		Upholsterer	Dorset Blandford	
		Harriet Dº	Wife	Mar		31		Hants Bishopstoke	
		Frederick Dº	Son		11		Scholar	Wilts Downtown	
		Caroline Dº	Daug			5	Dº	Dorset Blandford	
		Elizabeth Dº	Daug			2		Wilts Salesbury	
		George Dº	Son		1mo			Dº Dº	
		Reuben Down	Apprentice			15	Cabinet Makere Apprentice	Dº Fartierton	
75	High Street	Jonas Crawford	Head	Mar	36		Boatmaker	Dorset Ladco	
		Rebecca Dº	Wife	Mar		36	Binder	Middlesex London NW	
76	High Street	Francis Dixon	Head	Mar	30		Brushmaker	Wilts Salesbury	
		Mary Ann	Wife	Mar		27		Dº Dº	
		James Dº	Son			5		Dº Dº	

Scene from biblevideos.lds.org

MXF Challenges

- General Conference MXF containers package HD video, ASL video, up to 96 repeating audio tracks
- No known tool can validate file formats and extract metadata within an MXF container
- MediaInfo tool only extracts from MXF container
- Rosetta does not yet support ingest of repeating tracks

Scene from biblevideos.lds.org

MXF Solutions

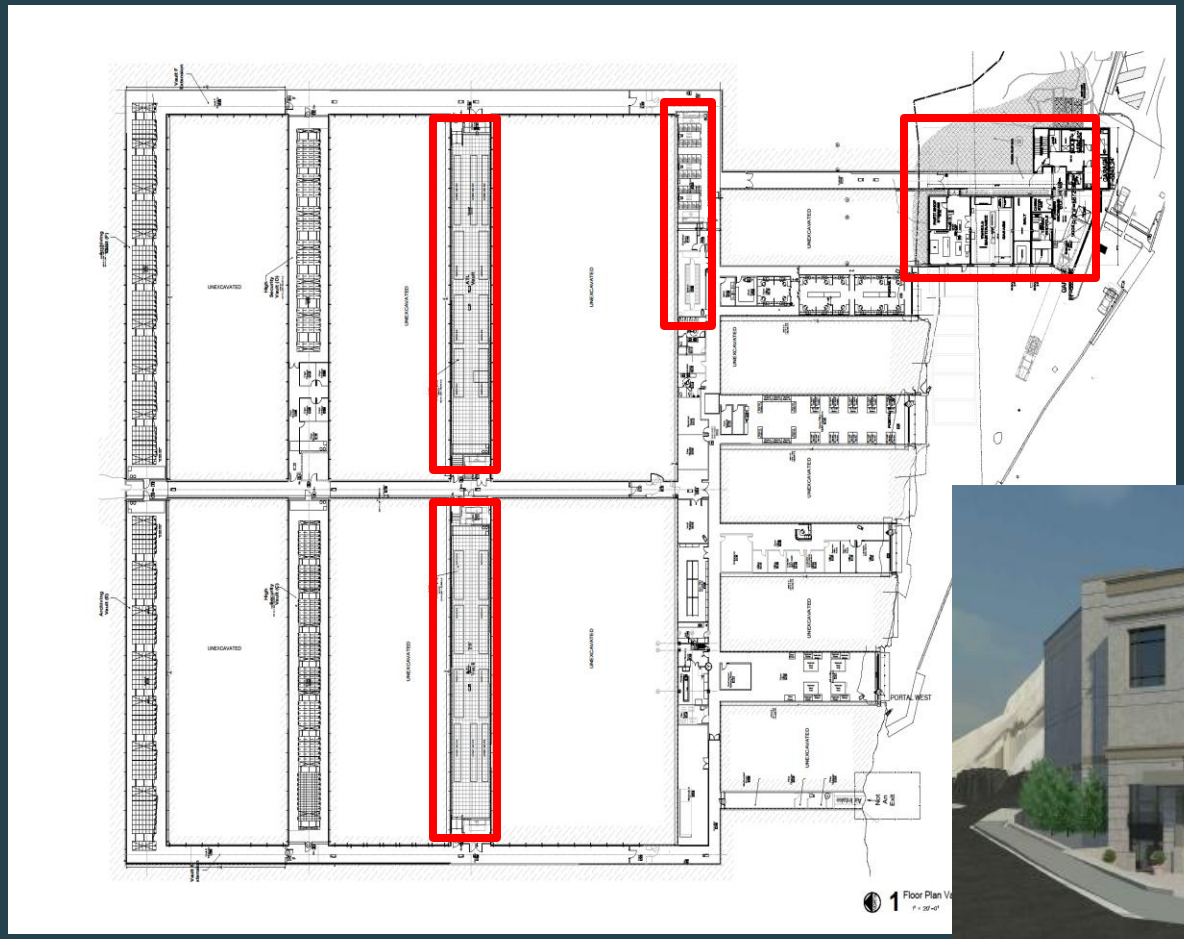
- Preservation team developed a Rosetta plug-in that utilizes MediaInfo to extract metadata
- MXF Extraction Tool concatenates data from repeating tracks in the metadata it extracts
- A modified MXF Extraction Tool given to Ex Libris (one video, one audio)
- *Available to all Rosetta users!*

Scene from biblevideos.lds.org

Granite Mountain Records Vault

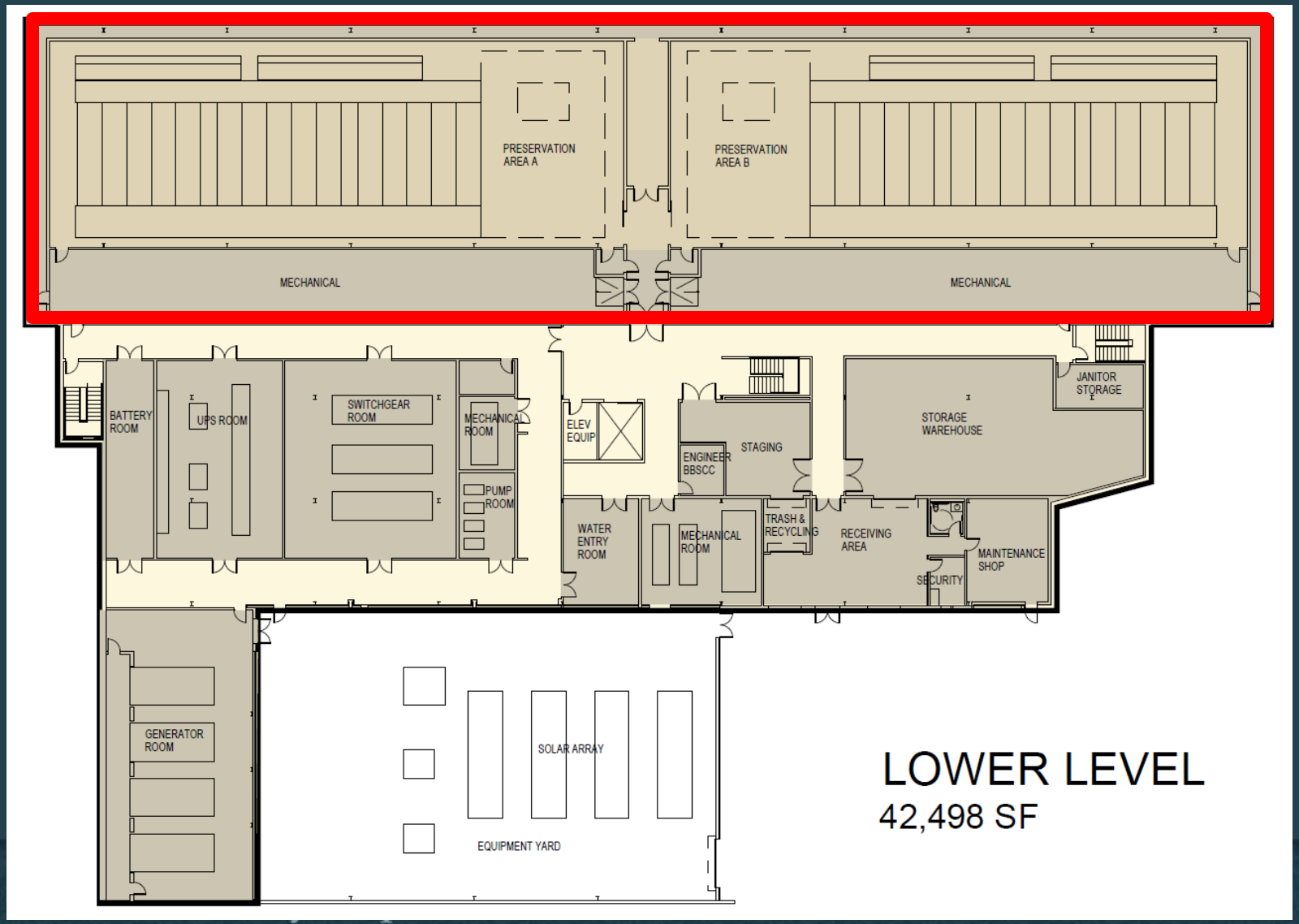
- Six tunnels bored into a solid granite mountain
- Stores FamilySearch microfilm collection and priceless artifacts
- Plans recently developed to renovate the facility for digital preservation

Granite Mountain Records Vault

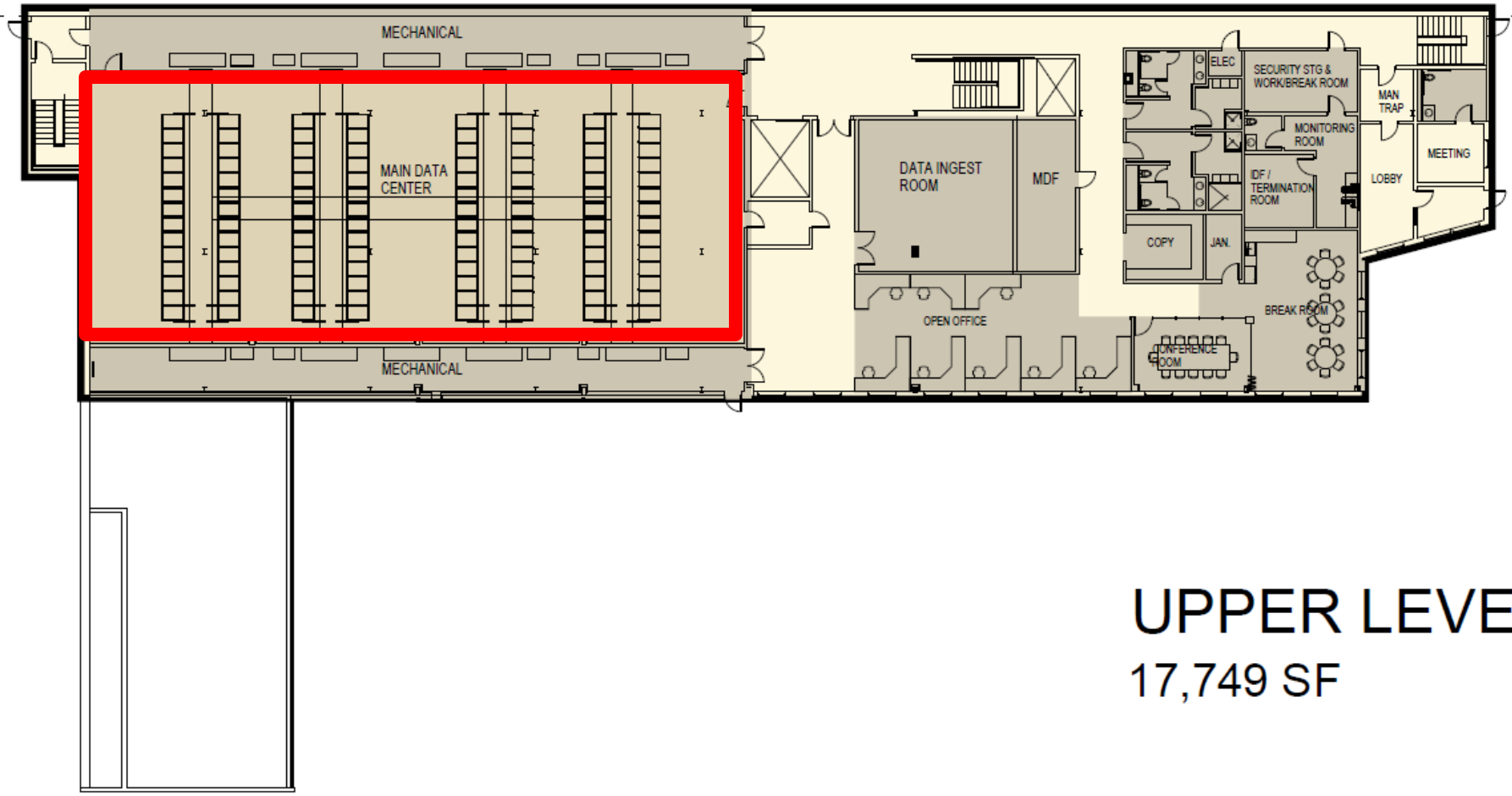


Remote Digital Preservation Facility

Remote Digital Preservation Facility



Remote Digital Preservation Facility



UPPER LEVEL

17,749 SF

Why Is the Church Doing All This?

- Build character
- Strengthen families
- *Foster personal and family happiness*

Thank you!

Questions?

Trademarks