Digital Preservation at The Church of Jesus Christ of Latter-day Saints

Gary T. Wright

wrightgt@ldschurch.org

IEEE MSST 2012

April 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

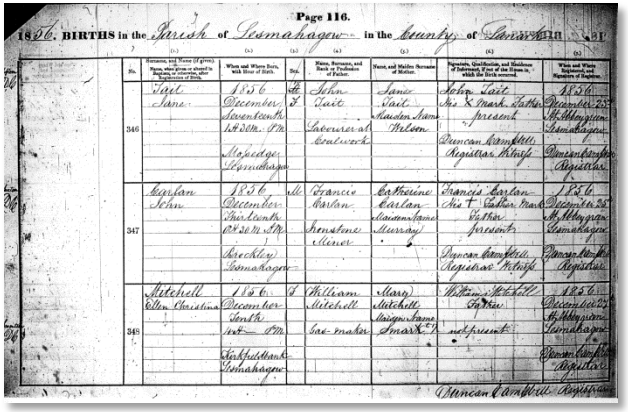
- Preserves records of enduring value from Church leaders, departments, universities, and affiliations (more than 35 organizations)
- Audiovisual records will consume vast majority of archive capacity
- 100+ PB in a decade for a single copy!

Church History Library on Temple Square

Family History Department

- Preserves the heritage of mankind
- 2.4 million rolls of microfilmed records
- 13.1 billion images of global historical records
- 300+ million images digitized annually
- 100+ PB in a decade!

FAMILYSEARCH



familysearch.org

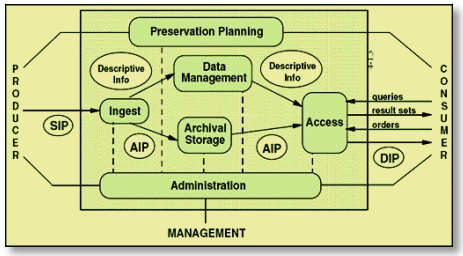
Architecting Two Preservation Systems

- CHD / FHD collaboration

- National Library of New Zealand business requirements

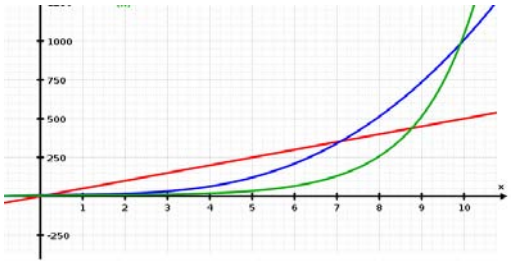
NATIONAL DIGITAL HERITAGE ARCHIVE PROGRAMME

- OAIS Reference Model



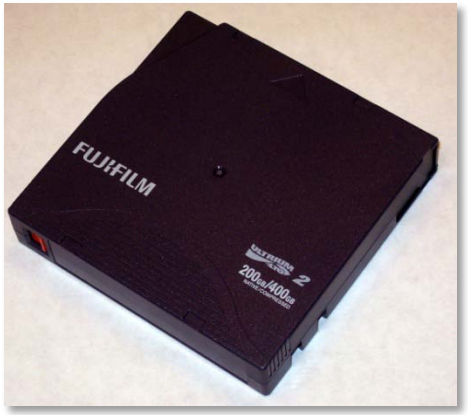
- Minimize cost of archival storage

- 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
- Tessella Safety Deposit Box SPOC and product feature comparison
- Conclusion—both products can meet Church needs

Different Business Needs

CHD

- Dark internal archive
- CHD ownership and service models
- Optimize for large audiovisual files
- Many different producers
- Many different file formats
- Each collection is a project
- Cataloging considerations

FHD

- Open public archive
- Offer software to partners
- Optimize for many small images
- One producer (pipeline)
- Only two file formats
- Everything is automated
- Faster need to scale

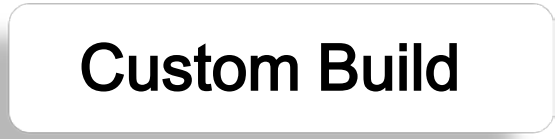
Storage Layers

CHD

ILM

Tape Interface

FHD



SDB Tape Adapter
(Single Tier)

CHD 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

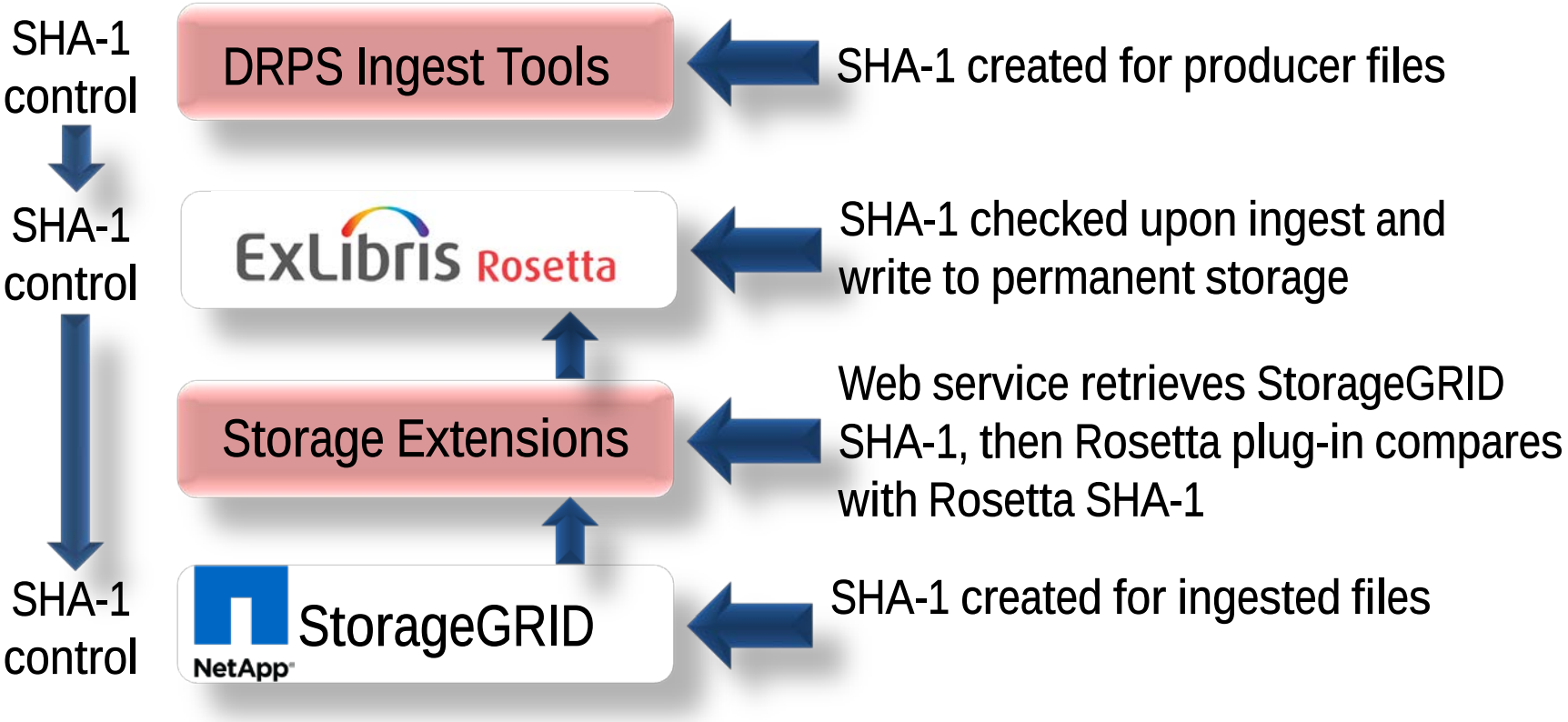
- A critical requirement of digital preservation—*differentiates an archive from other tape farms*
- Fixity information (secure hash checksum) enables data integrity validation by checking it every time data is written, transferred, moved, or copied
- End-to-end fixity checking should be performed from 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 information

Sizing Bit Errors

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

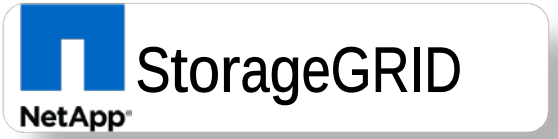
- 8 PB tape archive of videotaped interviews of Holocaust survivors and other witnesses
- Experienced 1500 bit flips in 8 PB (**2.3×10^{-14}** bit error rate)
- Real life measurements that provide guidance for tape archives

DRPS Data Integrity Validation



StorageGRID Fixity Checking

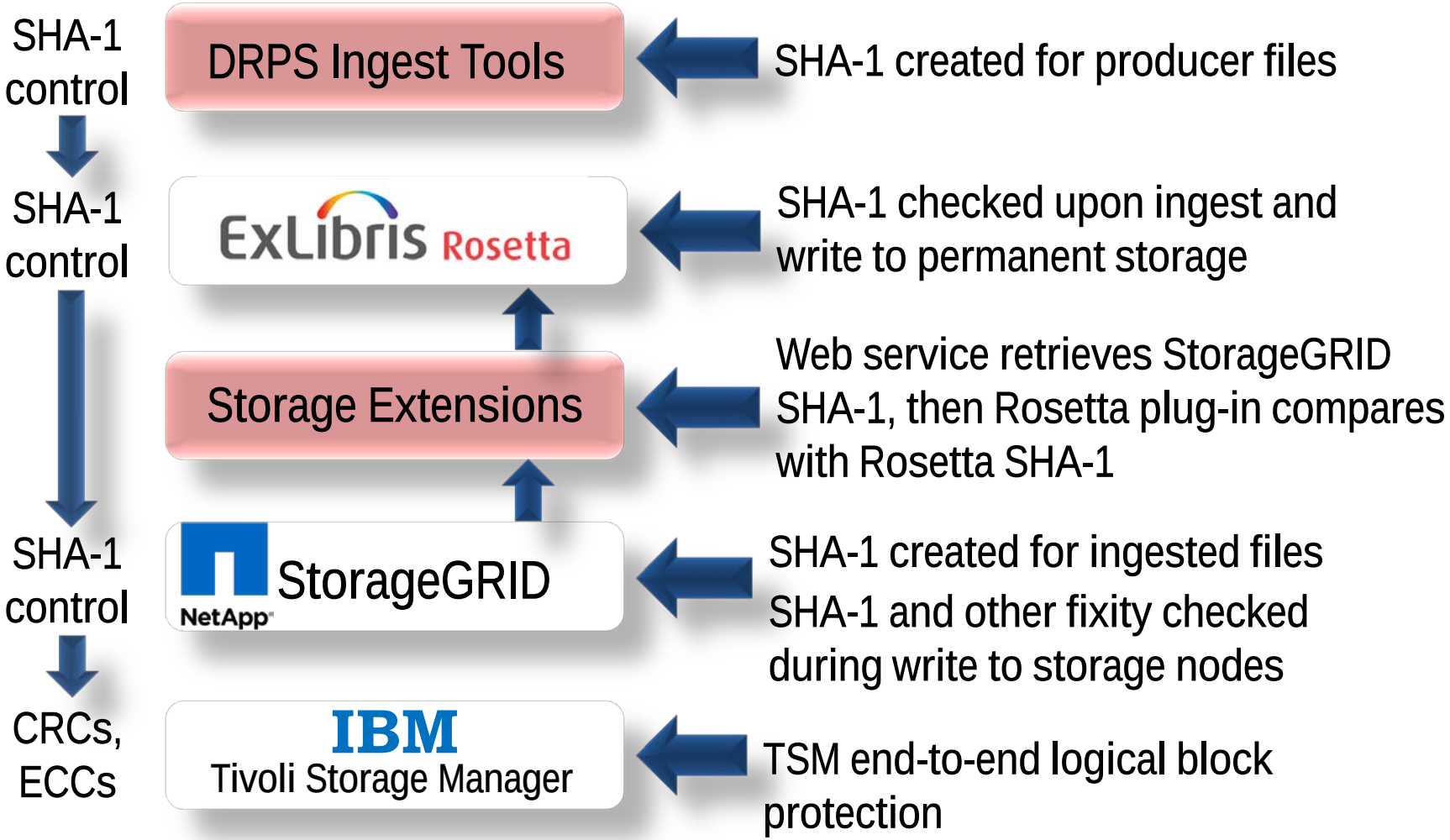
- StorageGRID is constructed around the concept of object storage



- 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. HMAC message authentication digest—checked on access

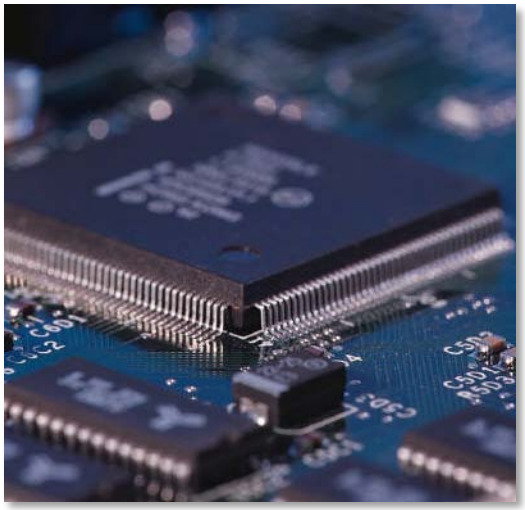
DRPS Data Integrity Validation



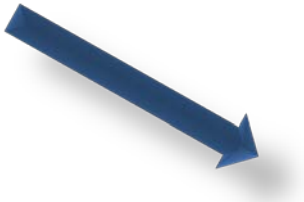
TSM End-to-End Logical Block Protection

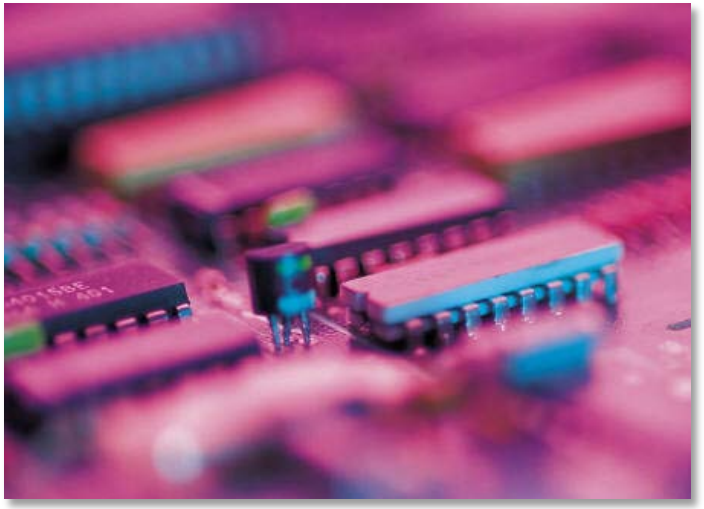
- Supersedes SHA-1 fixity information with cyclic redundancy check values (CRCs) and error-correcting codes (ECCs)
- Enabled with new, state-of-the-art functionality of IBM LTO-5 and TS1140 tape drives
- Seamlessly extends validation of data integrity until AIPs are correctly written to tape

TSM End-to-End Logical Block Protection



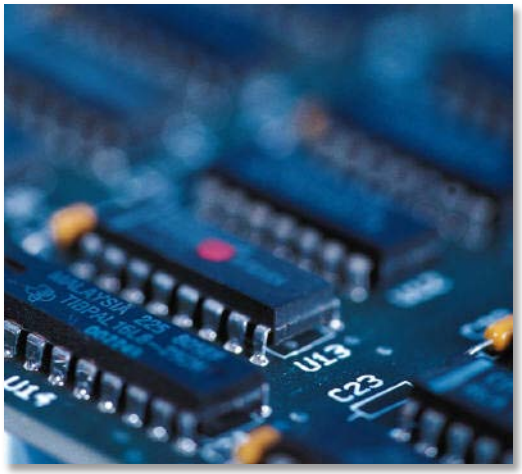
1. TSM server calculates and appends “original data CRC” to AIP logical block



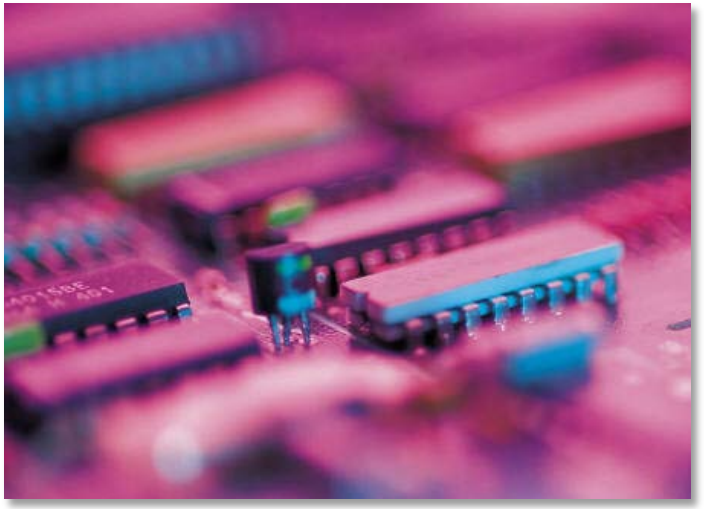


2. Tape drive computes its own CRC and compares to original data CRC

TSM End-to-End Logical Block Protection

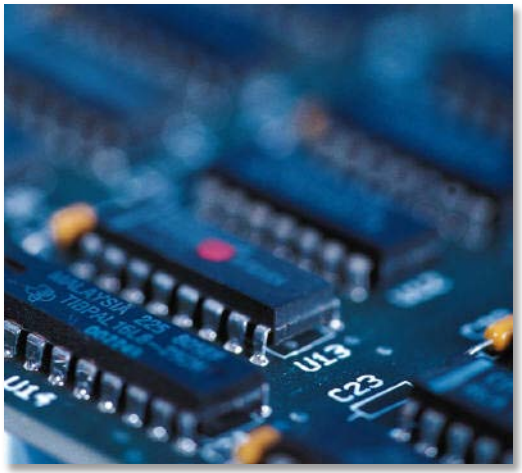


3. As logical block is loaded into drive data buffer, on-the-fly verifier checks original data CRC

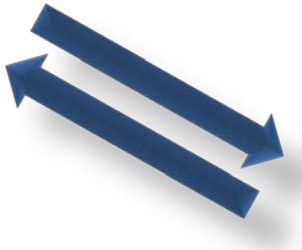


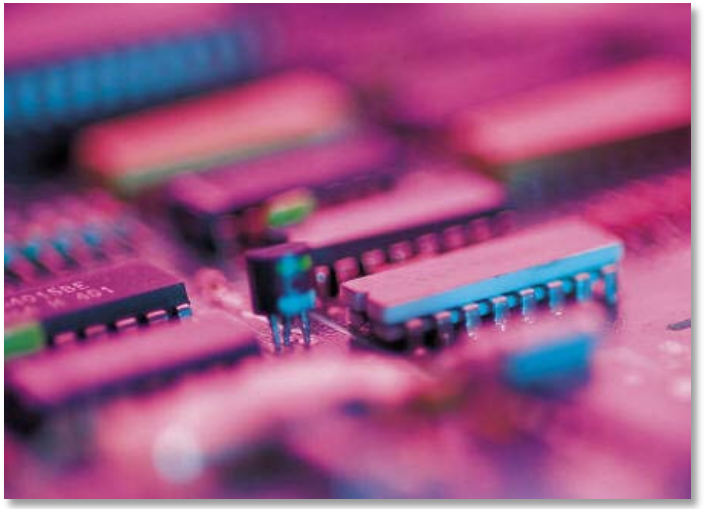
4. In parallel, a “C1 code” (Reed-Solomon ECC) is computed and appended

TSM End-to-End Logical Block Protection



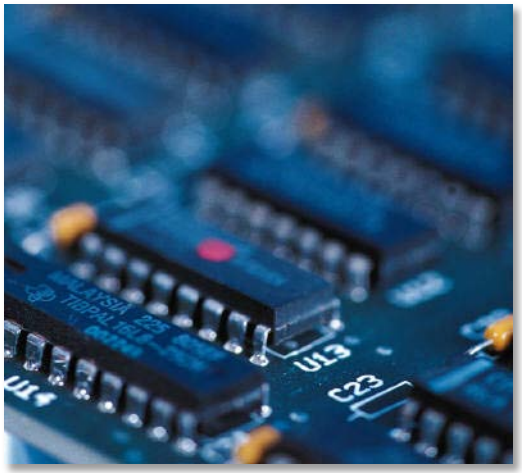
5. An additional ECC, referred to as “C2 code,” is added to the logical block



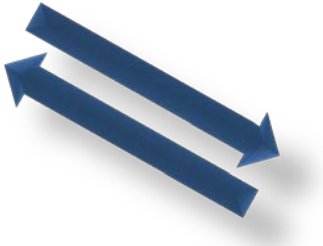


6. More powerful than the original data CRC, the C1 code is checked every time data is read from the buffer

TSM End-to-End Logical Block Protection

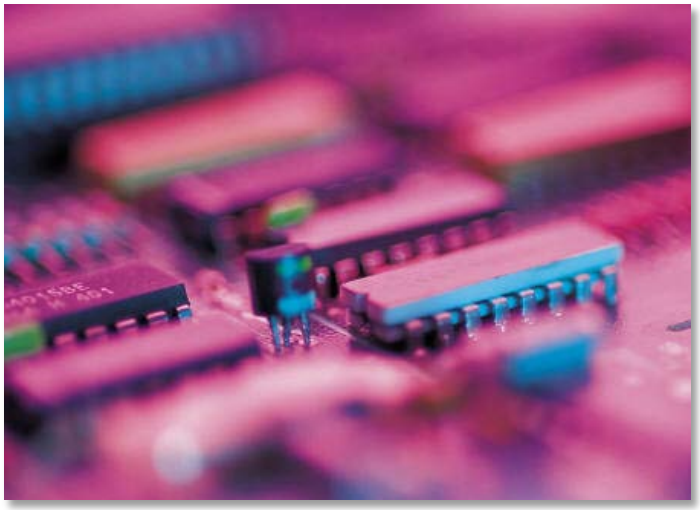


7. Data written to tape at full line speed with read-while-write process
8. Just written data loaded to buffer and C1 code checked



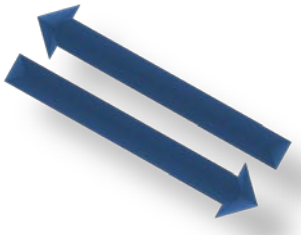
Successful read-while-write operation assures no data corruption from TSM client to tape

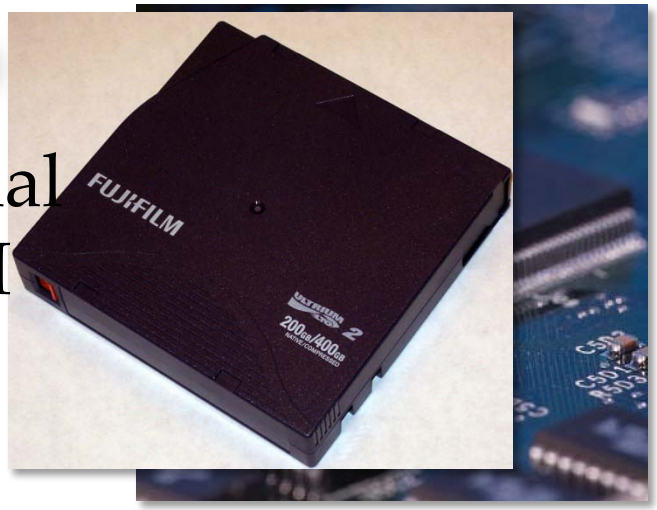
TSM End-to-End Logical Block Protection



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 original 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

10011011**0**100001011

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

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

image courtesy of IBM

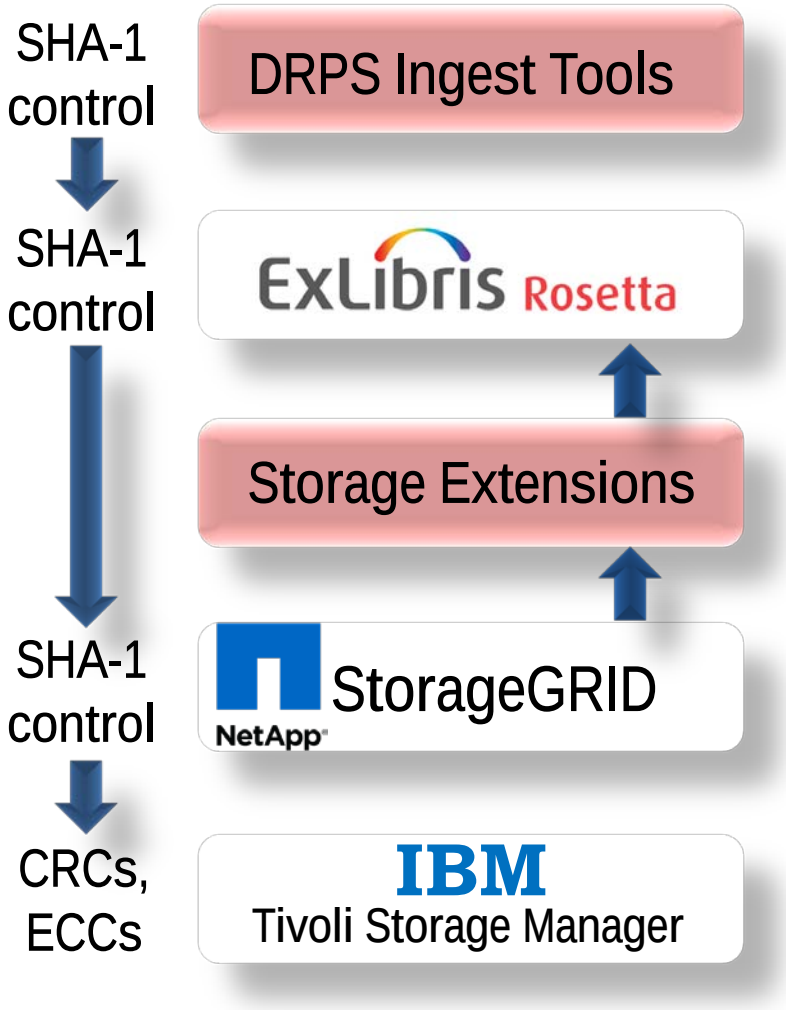
Ongoing Archive Data Integrity

- 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

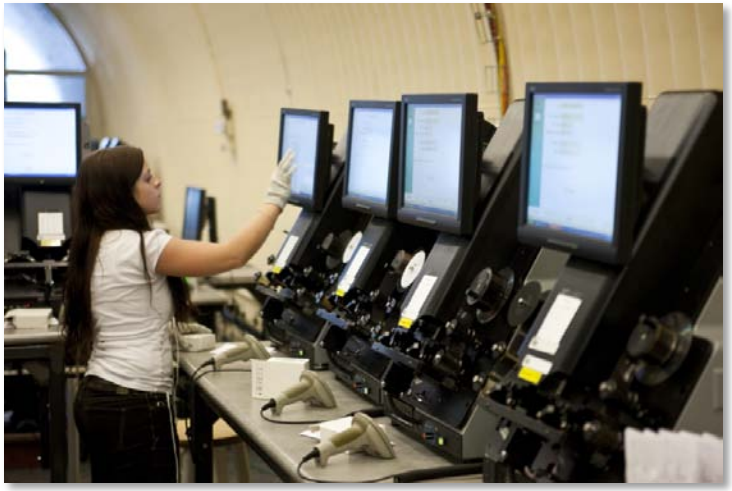


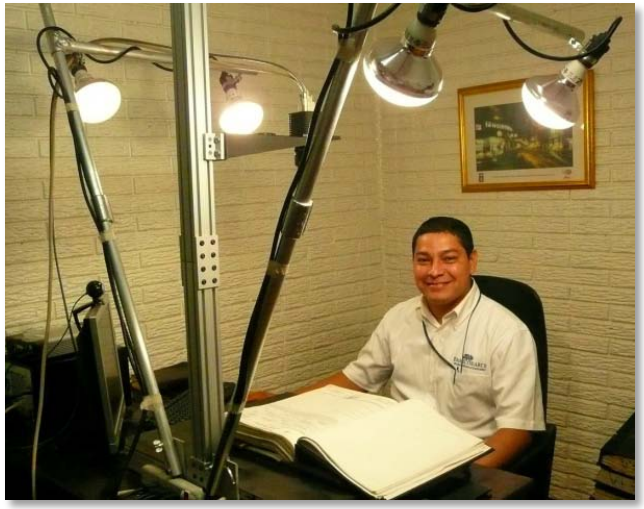
Optimizing the Use of Tape

- Rosetta was designed to work with NFS disk, so process automation does not optimize file access
- Rosetta automated processes are ordered by AIP
- Streaming media optimization requires that tapes be read and/or written **only once** to avoid tape thrashing (i.e., "shoe shining")
- "ordered by AIP" must become "ordered by file per tape"
- ICS, Ex Libris, NetApp, IBM collaborating on a solution

FamilySearch Digital Pipeline

- Processes 300+ million images annually
  - scanned microfilm
  - new digital photos
- Fixity information created at capture
- Low resolution images published on website, high resolution images preserved with DPS V3





FamilySearch DPS V3 Architecture

Digital Preservation System Version 3

Write Transfer Tapes

Pipeline Integration

Read Transfer Tapes

TapeWorkers

Preservation Functions

Tessella Technology & Consulting SDB

SDB Tape Adapter

TapeWriters

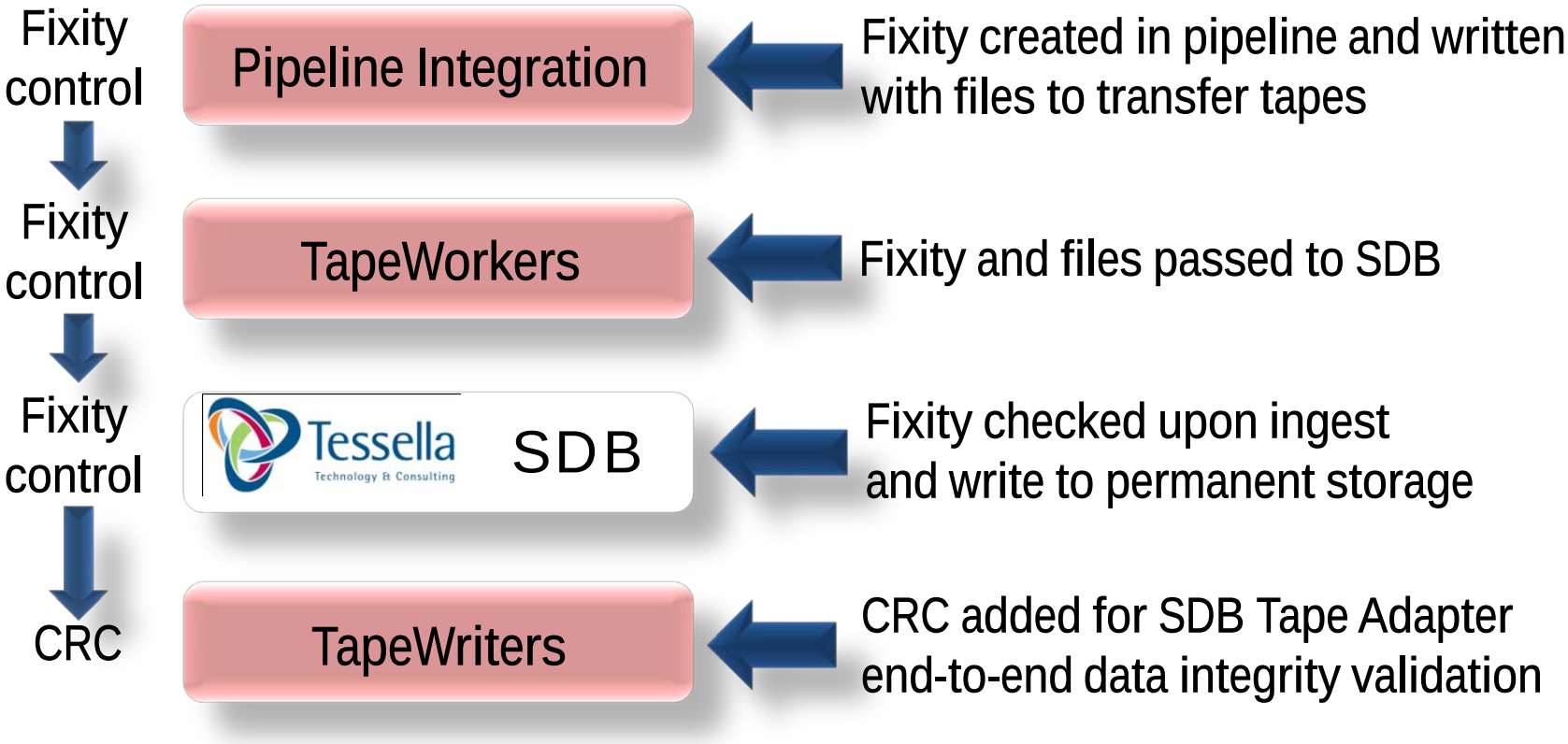
StorageTek T10000C Tape Drive

- SDB Tape Adapter end-to-end data integrity checking in Data Integrity Validation Mode
  - Tape Adapter appends CRC
  - Drive checks CRC before write
  - Drive appends CRC on read
  - Tape Adapter checks CRC

image courtesy of Oracle

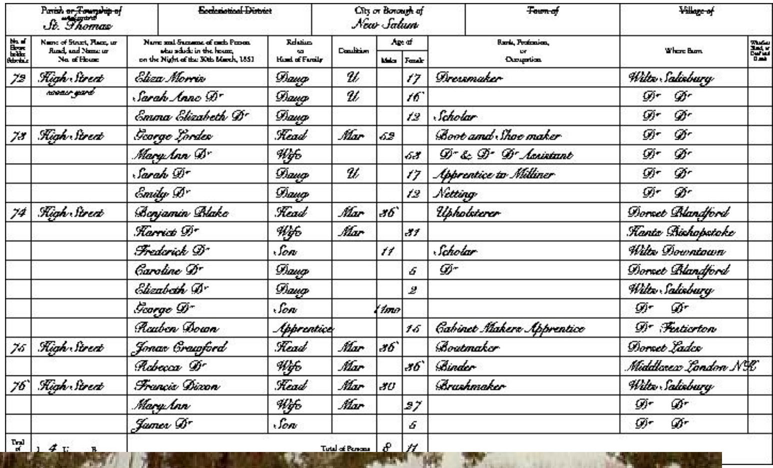
- In-drive Verify operation—a tape is mounted, drive checks CRC as data is being read (at full line speed)
- Similar to IBM drive capabilities

DPS V3 Data Integrity Validation



Preservation File Formats

- Both systems use lossless JPEG 2000 for still images
- FamilySearch is likely the world's largest user of JP2
- Motion JPEG 2000 wrapped in MXF containers used for audiovisual files preserved in DRPS



Parish ~~or Township~~ of St. Thomas		~~Ecclesiastical District~~		City or Borough of New Salum			~~Town of~~	~~Village of~~	
No. of House holder's Schedule	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	[illegible]
72	High Street (rear yard)	Eliza Morris	Daug	U		17	Dressmaker	Wilts Salisbury	
		Sarah Anne Dº	Daug	U		16		Dº Dº	
		Emma Elizabeth Dº	Daug			12	Scholar	Dº Dº	
73	High Street	George Lordes	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 Salisbury	
		George Dº	Son		1mo			Dº Dº	
		Reuben Down	Apprentice			15	Cabinet Makers Apprentice	Dº Fetterton	
75	High Street	Jonas Crawford	Head	Mar	36		Boatmaker	Dorset Lades	
		Rebecca Dº	Wife	Mar		36	Binder	Middlesex London NK	
76	High Street	Francis Dixon	Head	Mar	30		Brushmaker	Wilts Salisbury	
		Mary Ann	Wife	Mar		27		Dº Dº	
		James Dº	Son			5		Dº Dº	
Total of	4 U B	Total of Persons			8	11			

Scene from biblevideos.lds.org

MXF Challenges and Current Solution

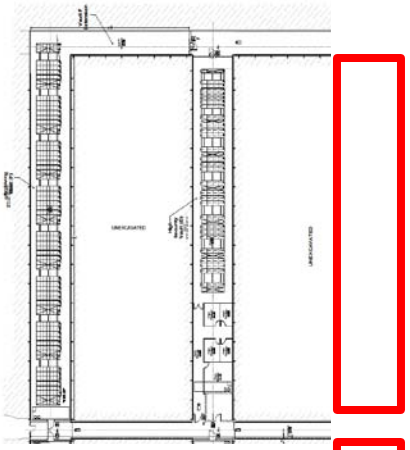
- No known tool can validate file formats and extract metadata within an MXF wrapper
- Open source MediaInfo tool only extracts metadata from the MXF wrapper
- ICS MXF Metadata Extraction Tool solves extraction problem
- Available to all Rosetta users

Scene from biblevideos.lds.org

Granite Mountain Records Vault

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

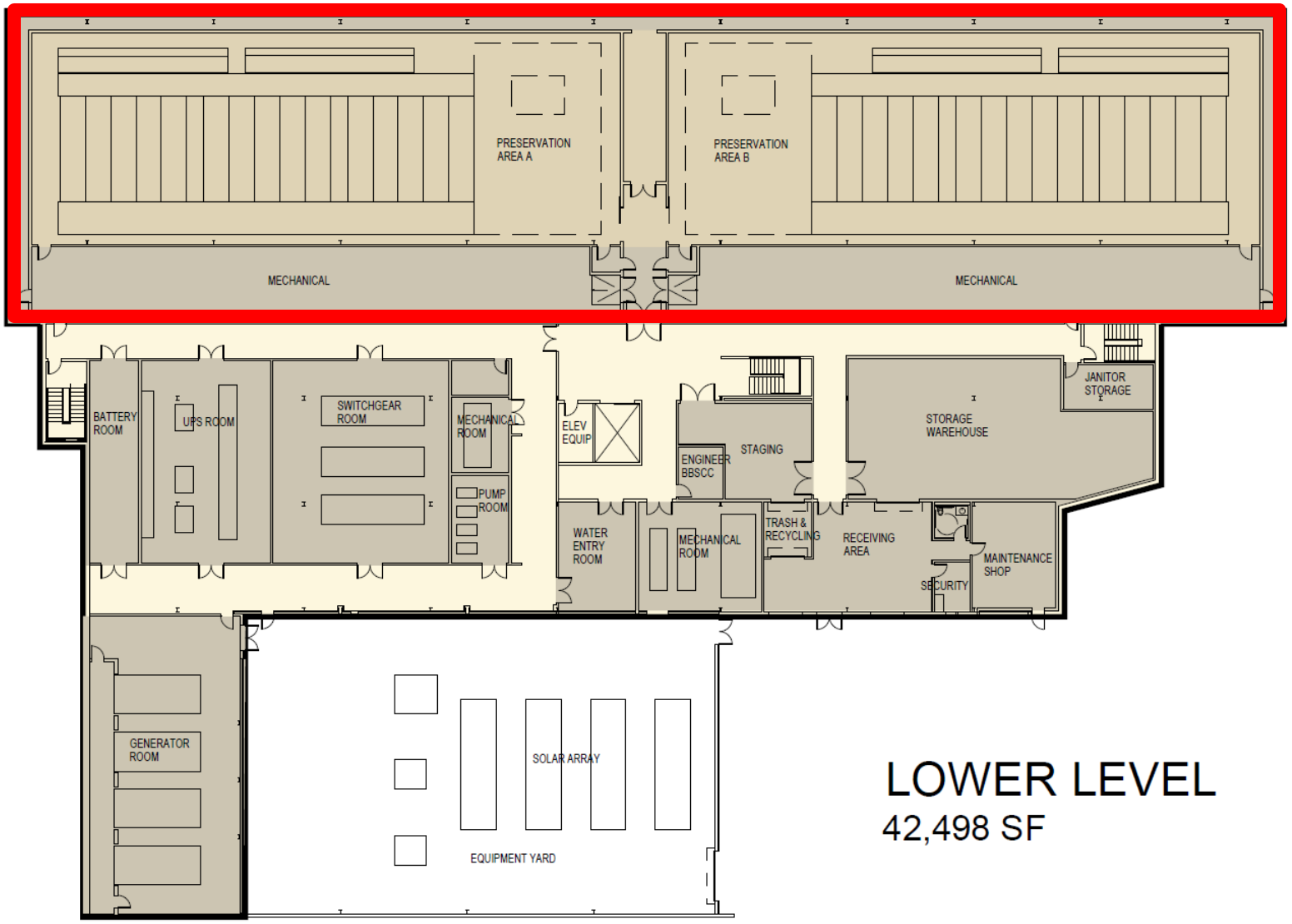
Granite Mountain Records Vault



Remote Digital Preservation Facility

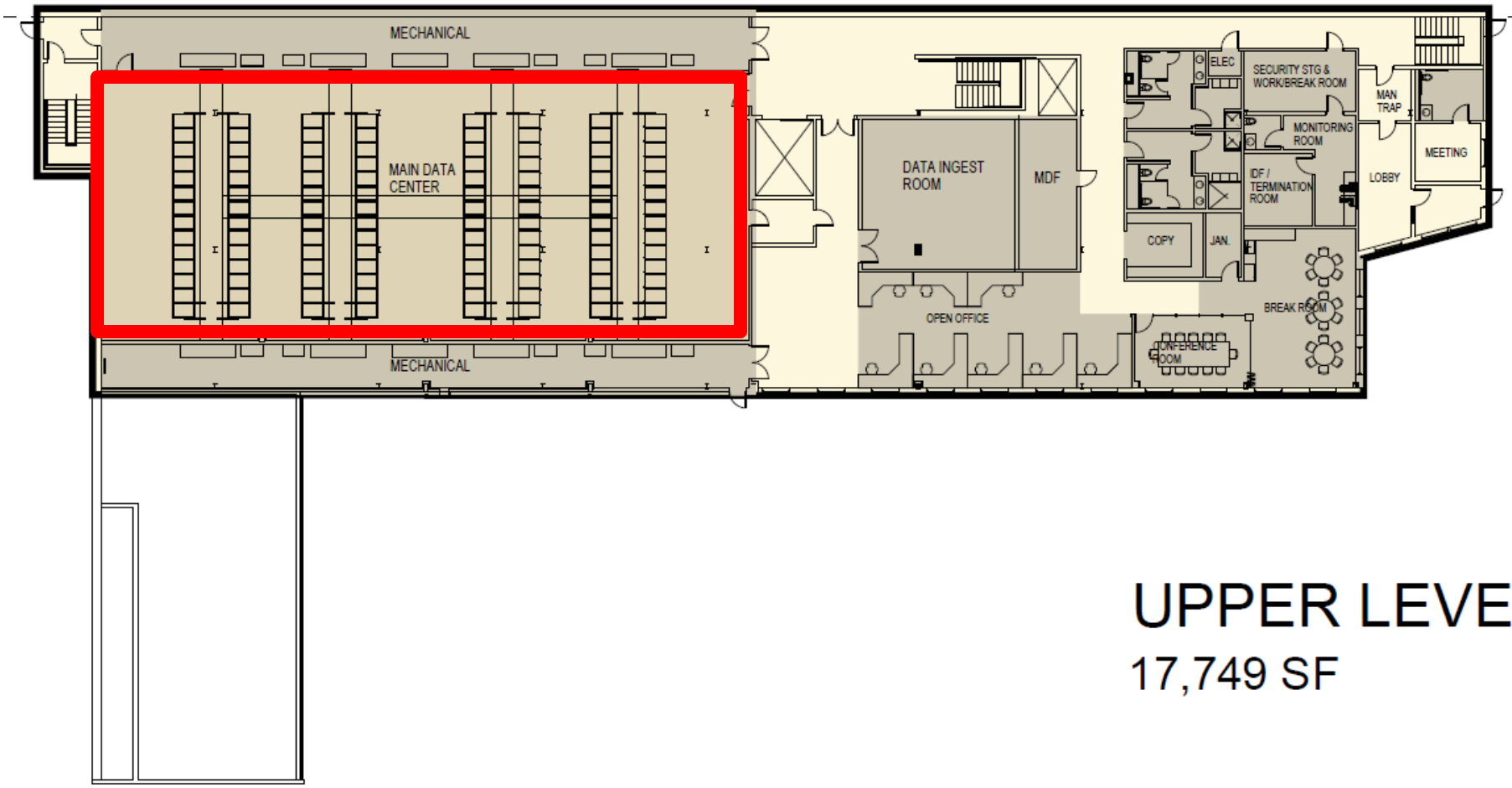
Design concept courtesy of Integrated Design Group

Remote Digital Preservation Facility



LOWER LEVEL
42,498 SF

Remote Digital Preservation Facility



UPPER LEVEL

17,749 SF

Why Is the Church Doing All This?

- Build character
- Strengthen families
- Link people of all races, religions, nationalities with their ancestors
- Preserve the heritage of mankind
- ***Foster personal and family happiness***

Thank you!

Trademarks