

OCIO

Office of the
Chief Information Officer
LIBRARY OF CONGRESS

Library of Congress Storage Architecture

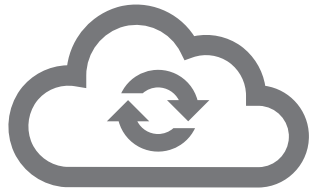
OCIO/ITSO/O&M
Operations & Maintenance



Overview

- Content Storage Environment Strategies
- Content Storage Environment Stats
- Content Storage Environment Projections
- Current Storage Architecture
- Perpetual Migration Continues
- Tech Refresh – ScoutAM Deployment
- Future – Converged Object Based Preservation Environment

Content Storage Environment Strategies

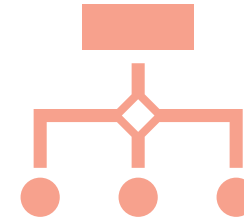


Preservation Deploying a 3:2:1 strategy

Three (3) copies of the data (2 tape copies and one cloud copy)

Two (2) technologies (Versity and AWS)

One (1) off premises copy (AWS Glacier)



Presentation Deploying a 2:2:1 strategy

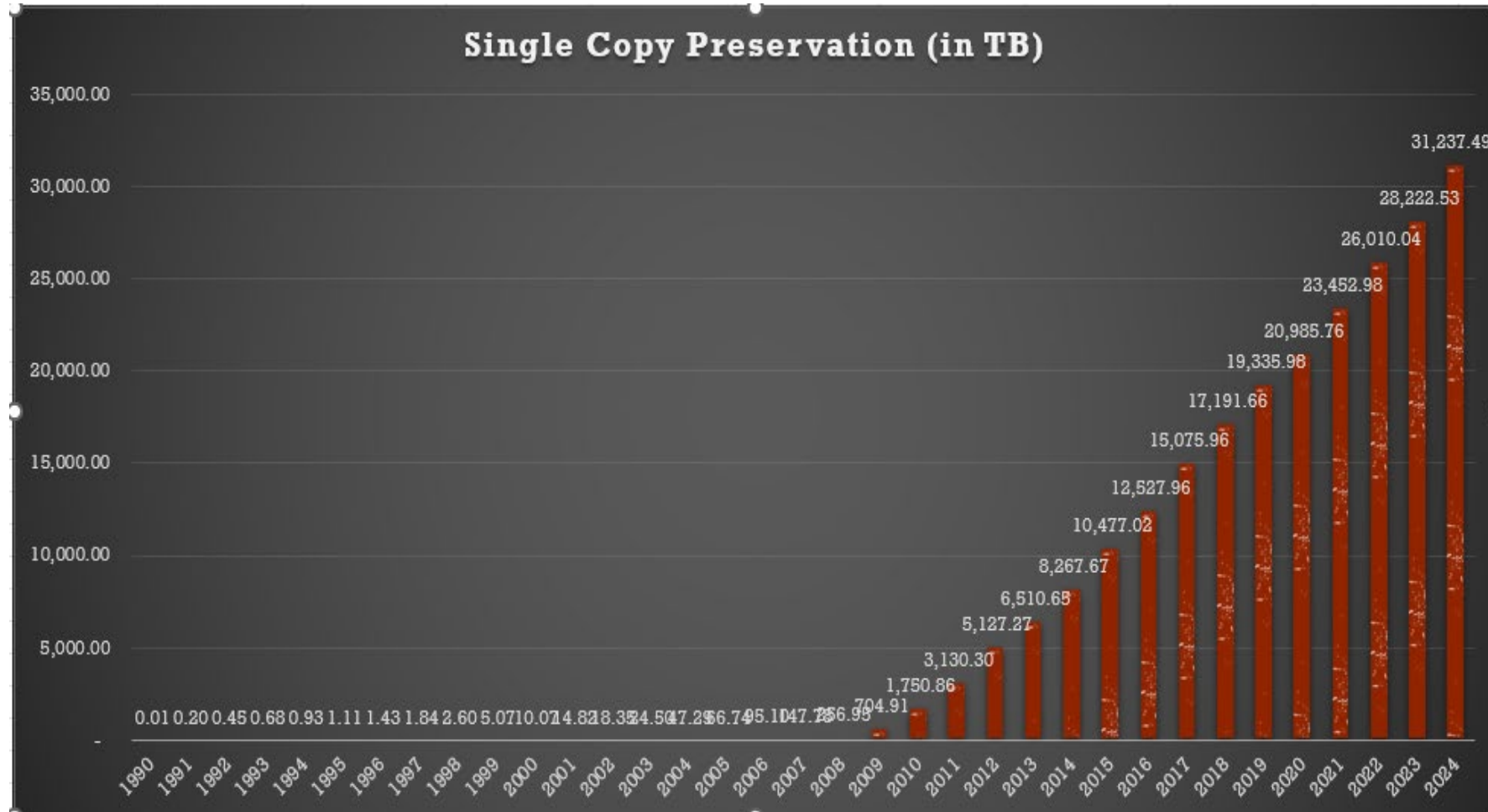
Two (2) copies of the data

Two (2) technologies (Spectrum Scale [gpfs] and AWS S3)

One (1) off premises copy (AWS S3 Intelligent Tiering)

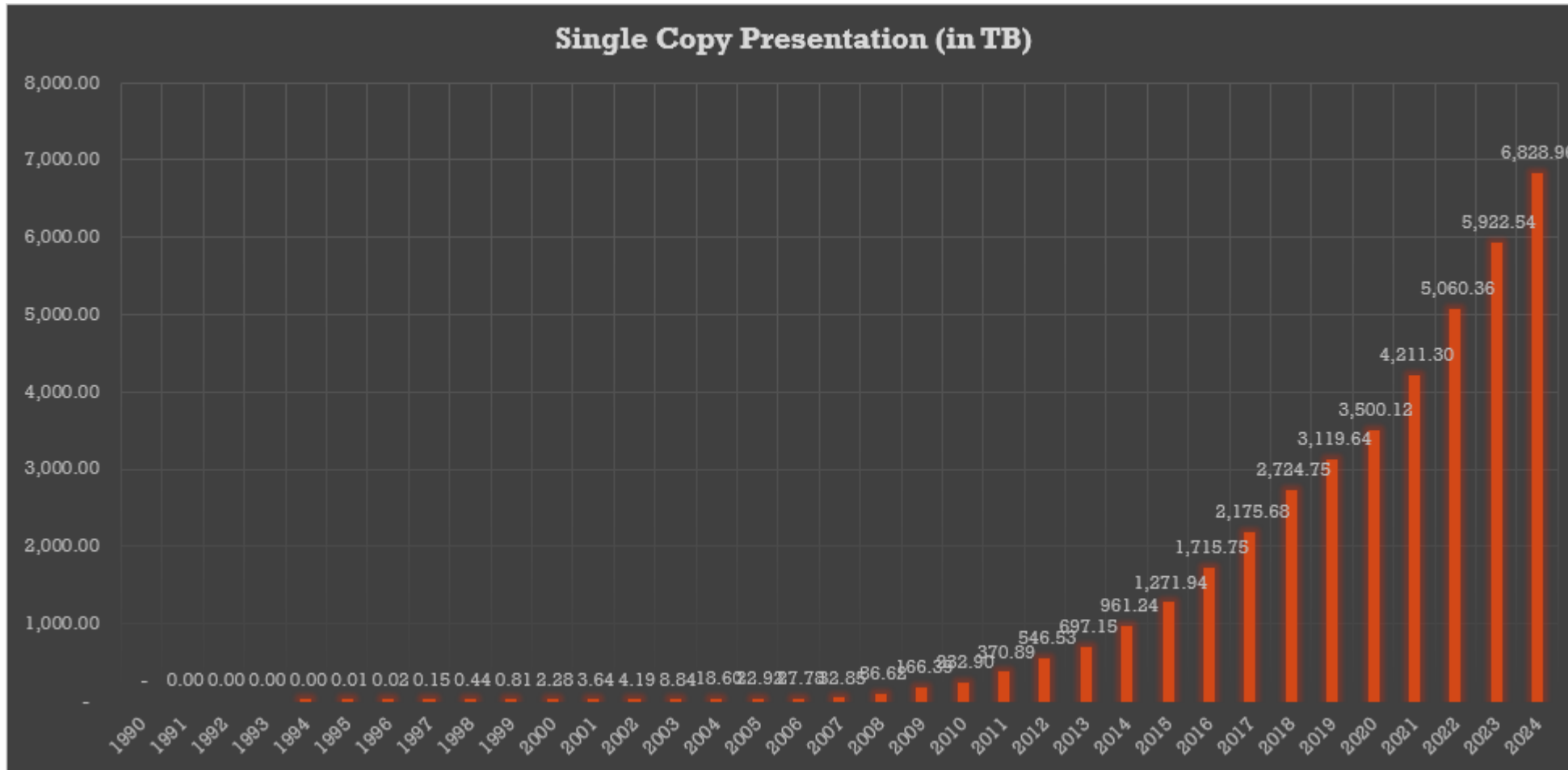
Content Storage Environment Statistics

- Preservation Single Copy 1.2B Objects (31.2PB) (10% Annual Growth)



Content Storage Environment Statistics

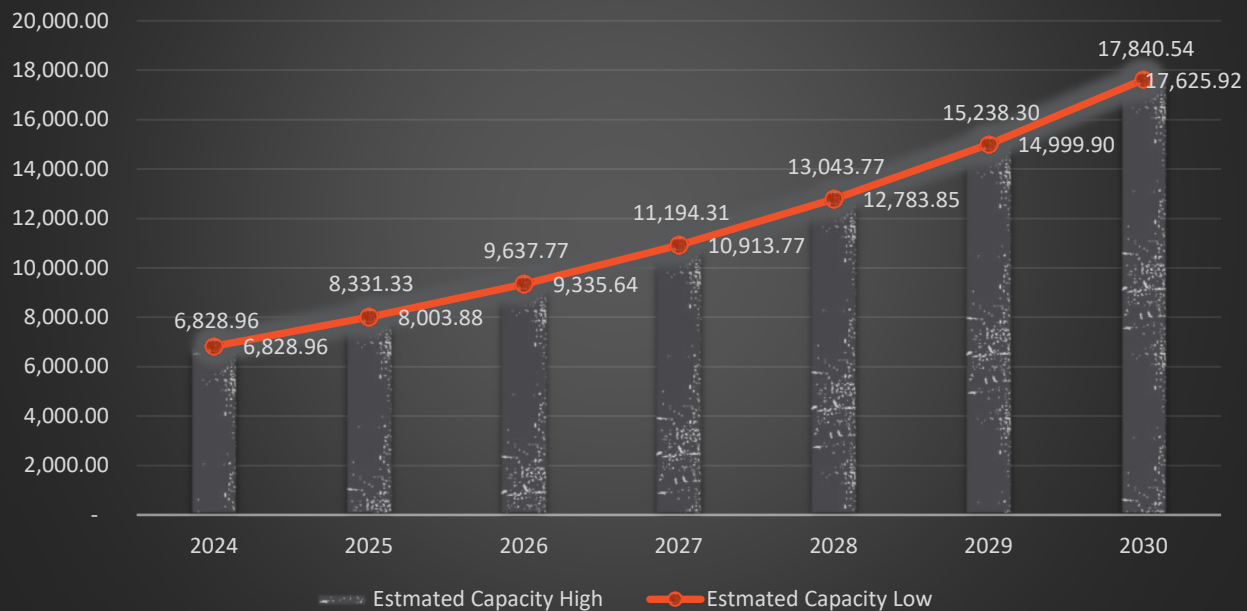
- Presentation Single Copy 288M Objects (6.8PB) (13% annual growth)



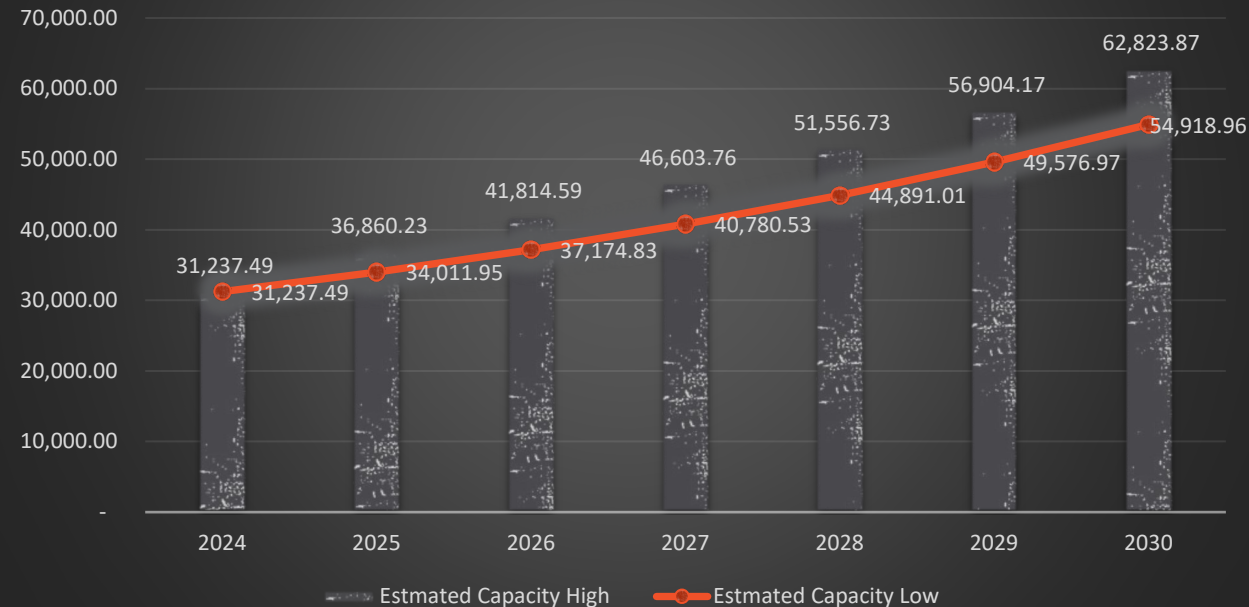
Content Storage Environment Projections

- Using a 10 and 5 year growth average
 - Presentation 19% to 14%
 - Preservation 13% to 9%

Presentation
Yearly Projected Growth
(Single Copy)

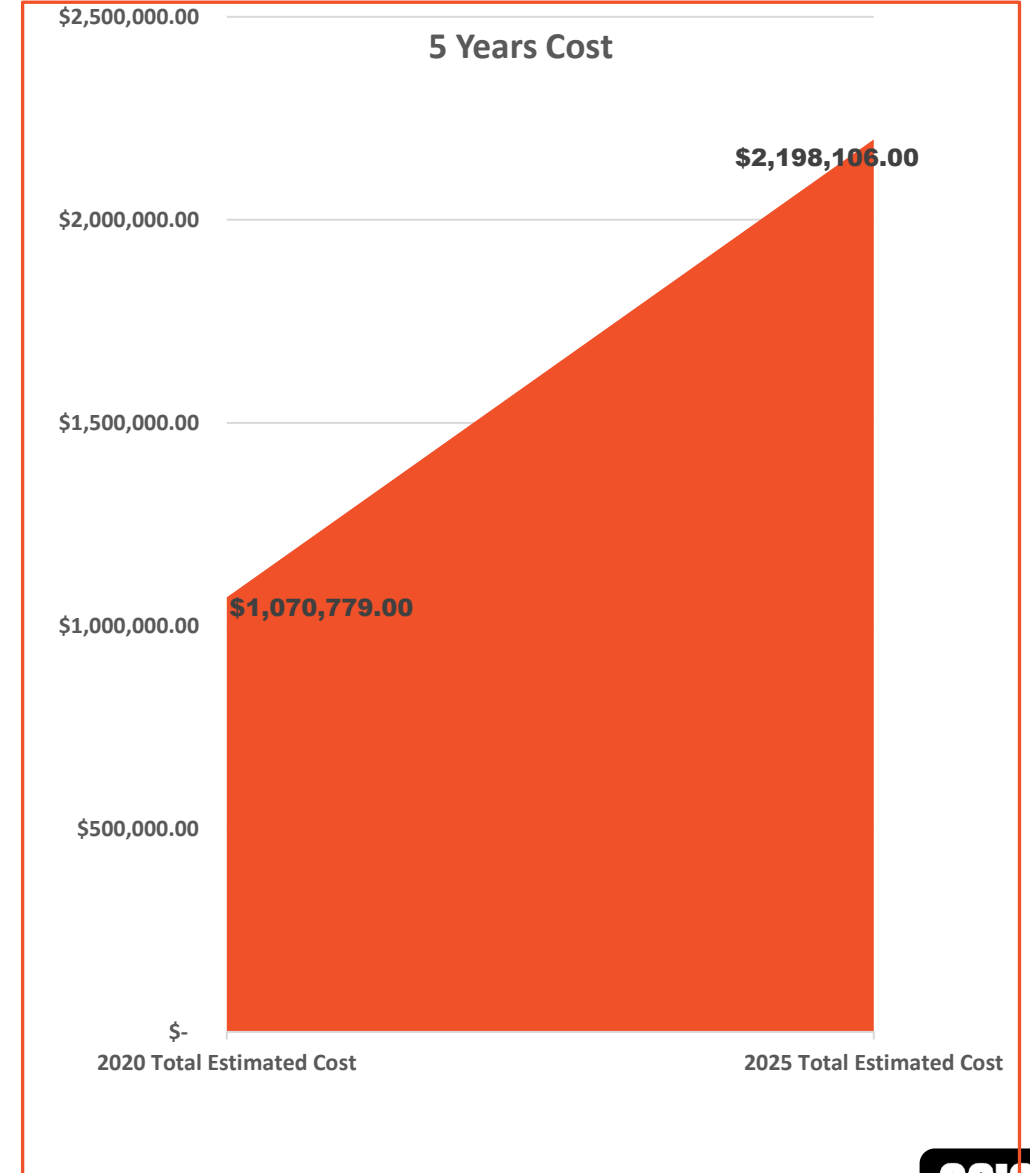


Preservation
Yearly Projected Growth
(Single Copy)

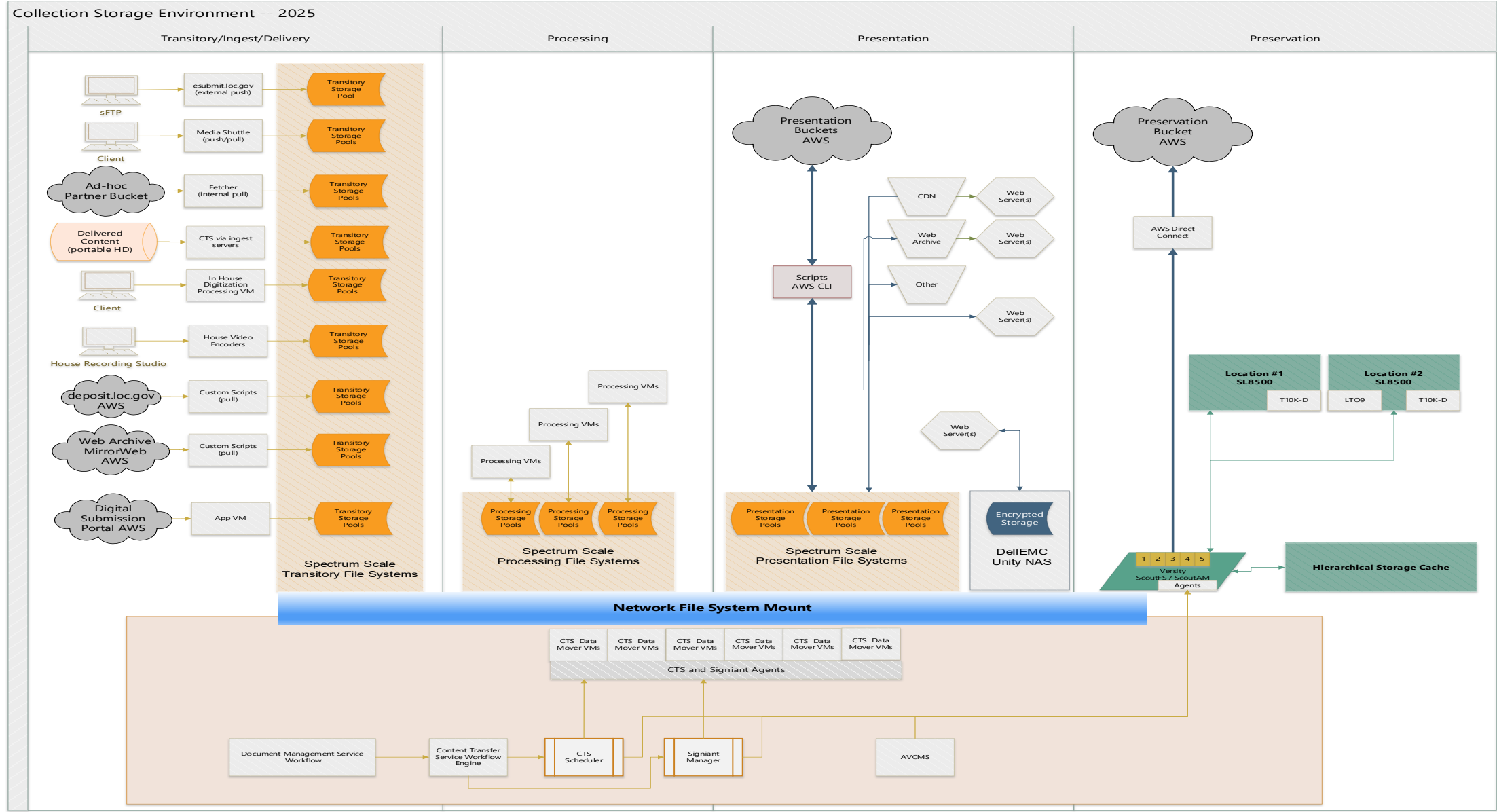


Looking at Overall Storage Cost

- Content storage has grown by 85% from 2020 to 2025 (increase from about 57PB to 105PB(48PB)) overall.
- Enterprise storage cost have increased by 105% in the past 5 years (estimated @\$1.07M/year to the current \$2.2M)
- 62.8% of the overall storage cost is associated with AWS S3 and Glacier adoption
- **COST DRIVERS**
 - Cloud storage for preservation (3rd) and presentation (2nd) copies
 - LOC digitization efforts (i.e. Film scanning, transcoding)
 - Born digital content submission (i.e. Web Archives, E-Submit, E-Deposit)



Current Storage Architecture



Perpetual Migration Continues

- Propagation of Preservation to AWS
 - Currently in year 6 of the project (97% complete)
- Software Technical Refresh
 - Versity Storage Manager to ScoutAM Archiver
- Upcoming Technology Change
 - Move to Converged Object Storage
 - Move to LTO9

Continuous Migration From the 2017 Conference
From the 2012 Conference



"Working on the cure for those suffering with digital Disposophobia"

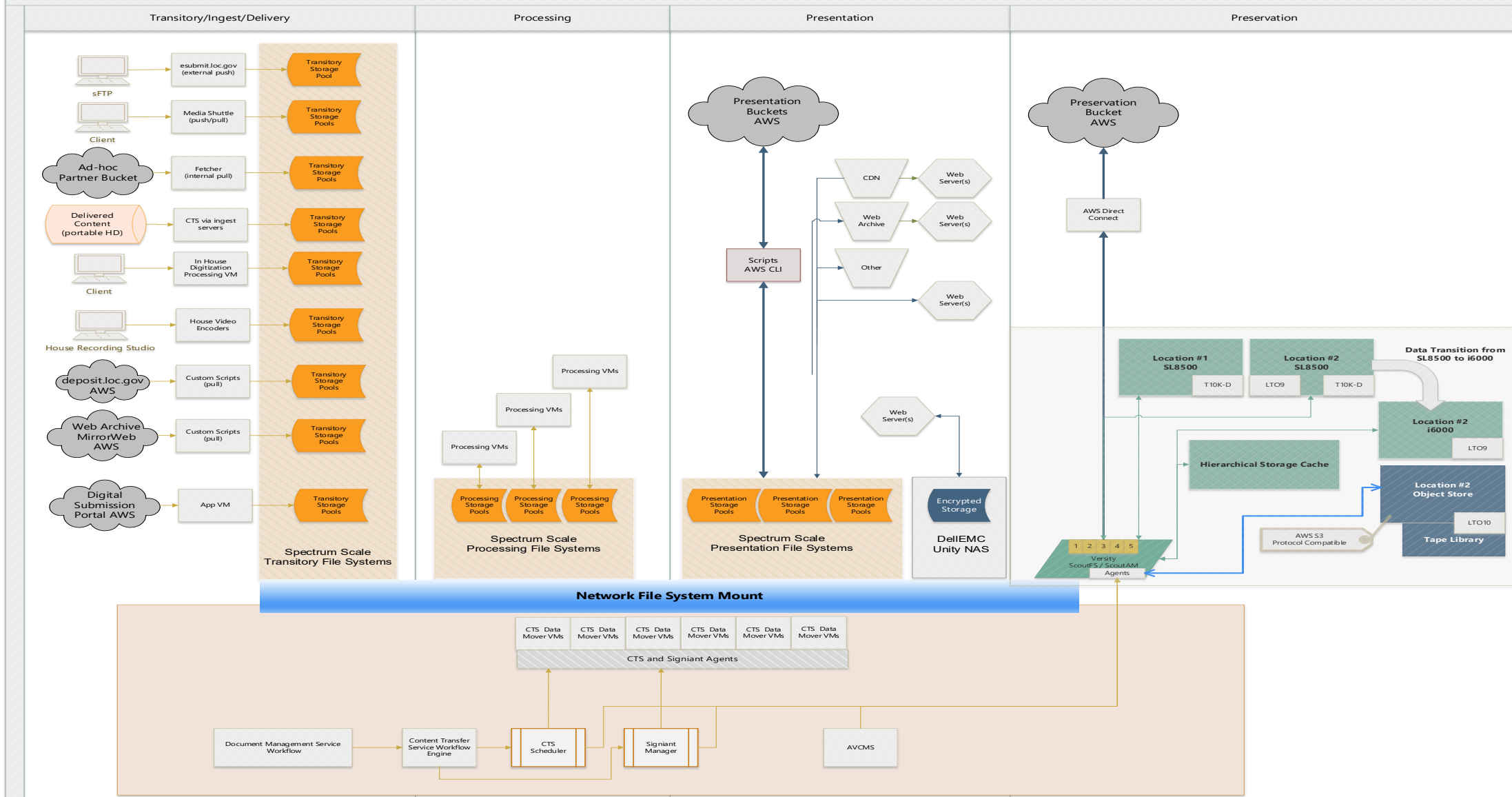
6

ScoutAM Deployment – Versity Tech Refresh

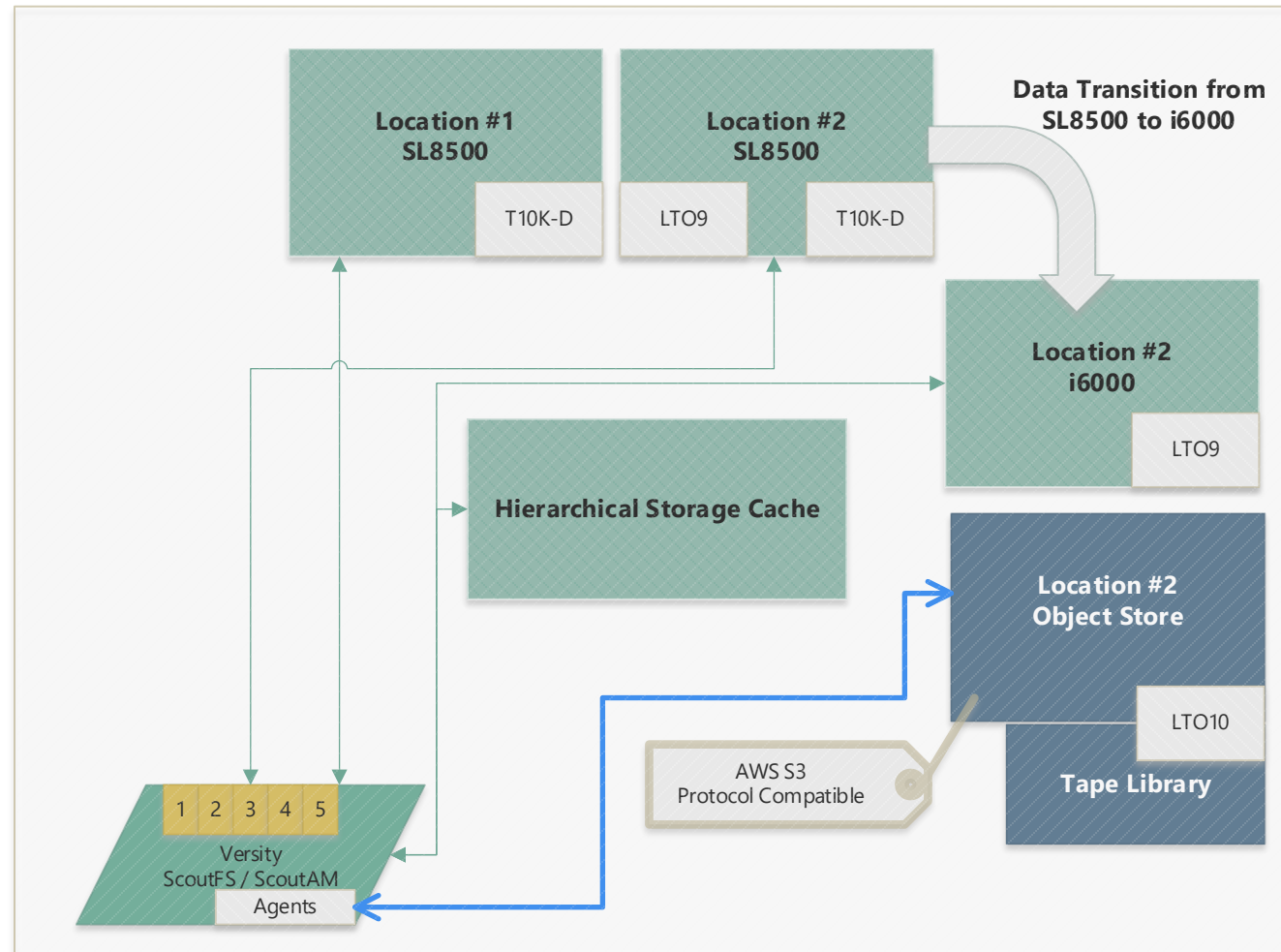
- Completed move from a single metadata server to a clustered model for availability and performance
- Upgrading Operating System from RHEL7 to RHEL8
- Moving one (1) copy from T10k-D to LTO9 tape drives
- Versity S3 gateway for object front-end to tape capability for testing
- Challenges thus far
 - Integration with legacy SL8500 tape libraries (ACSLs versus SCSI library control)
 - New interfaces and terminology versus familiarity with SAM-QFS

Future – Converged Object Base Preservation Environment

Collection Storage Environment -- 2025



Future – Converged Object Base Preservation Environment





Office of the
Chief Information Officer
LIBRARY OF CONGRESS

Thank you!

For More Information

Web: <https://digitalpreservation.gov/meetings/>

Points of Contact

Carl Watts, cwat@loc.gov

Chris Barr, cbar@loc.gov



Office of the
Chief Information Officer
LIBRARY OF CONGRESS

Questions?