

APT Trust

Academic Preservation Trust

Library of Congress
Designing Storage Architectures
March 27-28, 2022



ACADEMIC
PRESERVATION TRUST



Who are we?

A consortium of 18+ universities and
memory institutions across the US.





ACADEMIC
PRESERVATION TRUST

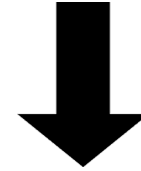
What do we do?



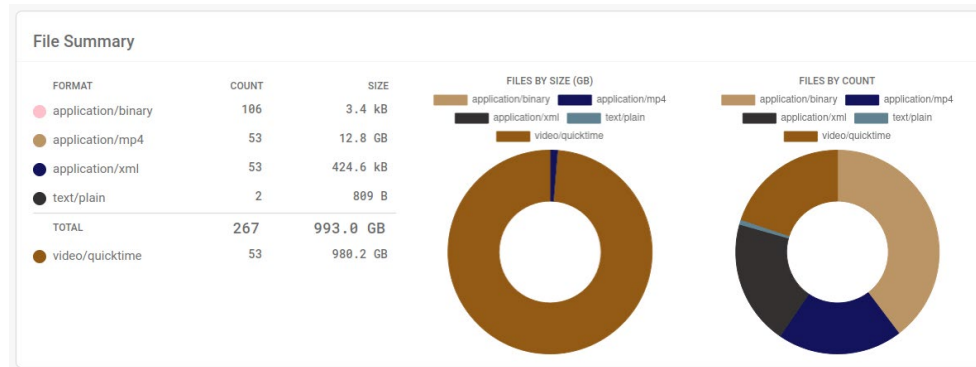
- Distributed digital preservation
 - Multiple geographic regions
 - Multiple storage technologies
 - Multiple cloud storage providers
- Regular fixity checks
- Access controls
- Auditing through PREMIS events
- Safe, multi-approval deletion workflows
- Large-scale data restorations
- Restoration spot tests
- Registry & REST API for managing deposited materials



Why us?



- Long-term preservation solutions are hard to build and maintain in-house.
 - Lack of expertise
 - Lack of political will and funding from above
 - Staff turnover & knowledge continuity
- Our members wanted a community-driven solution
 - Guided by & answering to depositors, not investors





What we built

- A stable, highly available, auto-scaling suite of services
 - Microservices running in Docker containers: Ingest, Restoration, Fixity, Deletion, Search & Discovery
 - Predictable performance under unpredictable loads
 - Cost effective
 - Requires near zero human intervention
- It took three tries and 8.5 years to get all of this right

Depositor

Ingest
Buckets

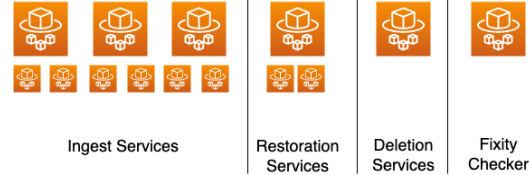


Depositor

Restoration
Buckets



Docker Containers in Fargate



AWS S3
Preservation



AWS Glacier
Preservation



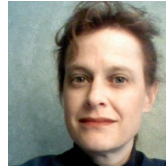
Wasabi S3
Preservation





Our Challenges

- We're a small team
 - One admin-security-ops lead
 - One developer



- Regular fixity checks aren't feasible for all storage solutions
 - Glacier
 - Glacier Deep Archive
 - Anything on tape or offline



- Cost Control
 - Architecting services for efficiency
 - Rooting out and taming hidden fees

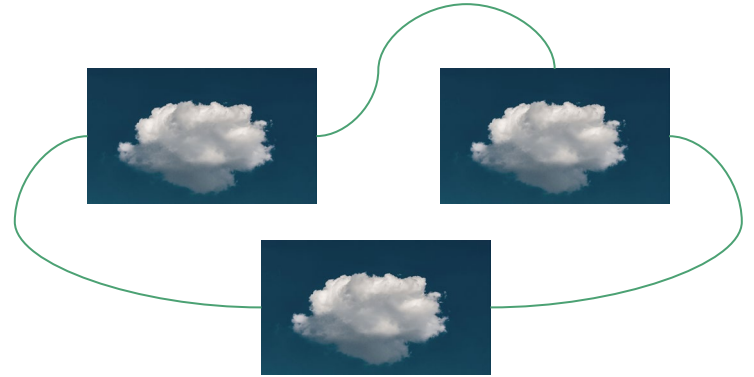




Our Challenges

- Multi-Cloud
 - Negotiate direct connects to reduce data transfer fees
 - Adding new vendors requires more legal/administrative work than tech work

- Data Privacy and Security
 - No uniform or permanent solution for PII: depositors must encrypt data before sending it
 - Identity and access management across a large suite of services
 - Minimizing our attack surface





ACADEMIC
PRESERVATION TRUST

Want to know more?

Website: <https://aptrust.org/>

Contact: help@aptrust.org

Wiki: https://wiki.aptrust.org/Main_Page

How To: <https://aptrust.github.io/userguide/>



Image Credits

Disk Drive	Photo by Patrick Lindenberg on Unsplash
Card Catalog	Photo by jesse orrico on Unsplash
Guy Fawkes Mask	Photo by Adnan Khan on Unsplash
Padlock on Computer	Photo by FLY:D on Unsplash
Cloud	Photo by C Dustin on Unsplash
Glacier	Photo by Cassie Matias on Unsplash
Tape Drive	Photo by Scott Beale on Flickr
Burning Cash	Photo by Jp Valery on Unsplash