DURACLOUD™

Store, Manage, and Archive Content in the Cloud

Michele Kimpton, DuraSpace
CEO DuraSpace
Nate Klingenstein, Internet 2
Internet 2 meeting, April 2013

Timeline

- April 2012 began discussions

- Sept 2012 signed agreement

- Oct 2012 started validation process with:
  - U of Virginia
  - North Carolina State University
  - U of Michigan

- April 2013 service validation complete-GA
  - Shibboleth integration, Incommon authentication
  - Business Agreement
  - Validate offering

What is DuraCloud?

Archiving and preservation services in the cloud

Ability to choose one or multiple cloud storage providers
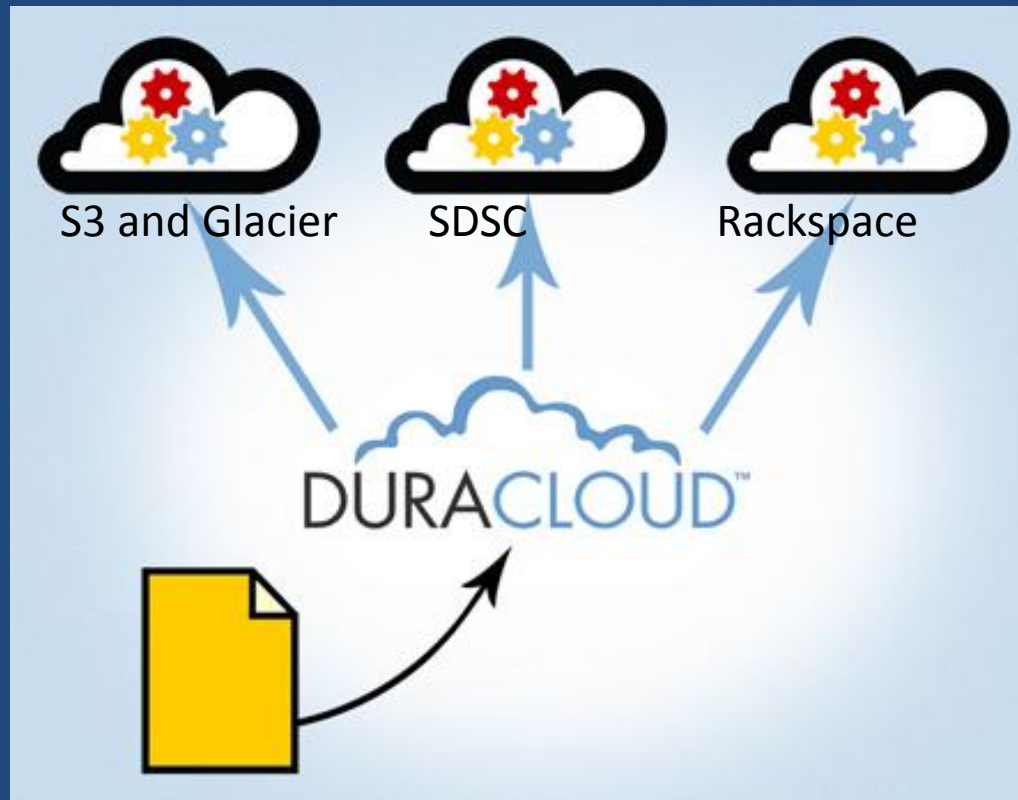
What does DuraCloud do?

Online backups

File health checking

File synchronization and healing

Online sharing

Media streaming

Image serving

How can DuraCloud help me with my archiving and preservation strategy?

- ✓ Ability to easily move online copies offsite, distribute geographically

- ✓ Easily check the health of all copies

- ✓ Easily repair or replace corrupted files

- ✓ Easily move or download content from multiple providers

- ✓ Scale storage up or down as needed

- ✓ Full Web access with ability to view/edit/delete and upload/download your content

Why should you care?

- Government mandates in place to provide long term access to content well after the grant period

- Contractual obligations for managing content

- Ever increasing amounts of digital content and digital data that needs to be archived

Cost

- $100-$1,000 per TB/yr

- Annual basic subscription fee $500/yr

- No I/O cost

- No data transfer costs

- Can pay as you go, or annual contract to lock in cost

Report by MHS, April 2013

http://www.mnhs.org/preserve/records/specialprojects.htm

The following table summarizes the capabilities and cost information for the various vendors.

Provider / Service	Data Integrity	Reliability	Scalability	Retention and Portability	Availability	Data Ownership	Preservation Functionality	Total 3-Yr. Costs
Amazon S3	Limited checksums	Average	Almost unlimited	Not easy to move	Average	Similar to others	None	Medium
Amazon Glacier	Limited checksums	Multiple tape copies	Almost unlimited	Not easy to move	Lower since on tape	Similar to others	None	Low
Google Cloud Storage	No checksums	Average	Almost unlimited	Not easy to move	None in contract	Contract concerns	None	Medium
Tessella's Preservica	Checksums and CRC	Multiple cloud copies	Same as S3	Multiple providers	Multiple cloud copies	Similar to others	Developed for this	?
VISI ReliaCloud	Limited checksums	Average	Cannot support MHS	Somewhat limited	Average	Similar to others	Some claimed	High
SDSC Cloud Storage	Automatic verification	Average but no tapes	Almost unlimited	Claimed to be easy	Concerns about disks	Similar to others	None	Medium
DuraSpace DuraCloud	User run checksums	Multiple cloud copies	Same as S3	Multiple providers	Same as S3	Similar to others	Some claimed	Medium
IBM SmartCloud	Limited checksums	Average	Unknown	Unknown	None in contract	Similar to others	None	?
FujiFilm Permivault	Custom plans	On-site and cloud	Almost unlimited	Somewhat limited	On-site copy	Similar to others	Limited	Low
FujiFilm Permivault Client	Custom plans	Cloud only	Almost unlimited	Somewhat limited	Average	Similar to others	Limited	Low
Code 42 CrashPlan Pro	Limited checksums	Average	More limited than S3	Somewhat limited	Average	Similar to others	None	Low

Table 4 – Comparison of Capabilities and Pricing [15]

The legend below defines the colors used in the table above:

(green)	Very good to best capability
(yellow)	Average capability
(red)	Poor or no capability
(orange)	Unknown data

What users are saying

- "The ***ease-of-use*** of DuraCloud is its most impressive feature." ~Columbia

- "DuraCloud delivers the ***benefits*** of a diverse network of storage locations, but ***without the overhead*** of managing different vendors." ~ICPSR

- DuraCloud provides a robust infrastructure to support digital preservation of content in our institutional repository, along with services that enhance the delivery and integrity of the content."

  ~Geneva Henry, Rice University

In production for one year

- Some stats-
  - 35 TB of data
  - 25+ institutions using Duracloud as a service
  - 100% Durability, no file failures
  - 5.46 million items stored
  - 60+ identified file types
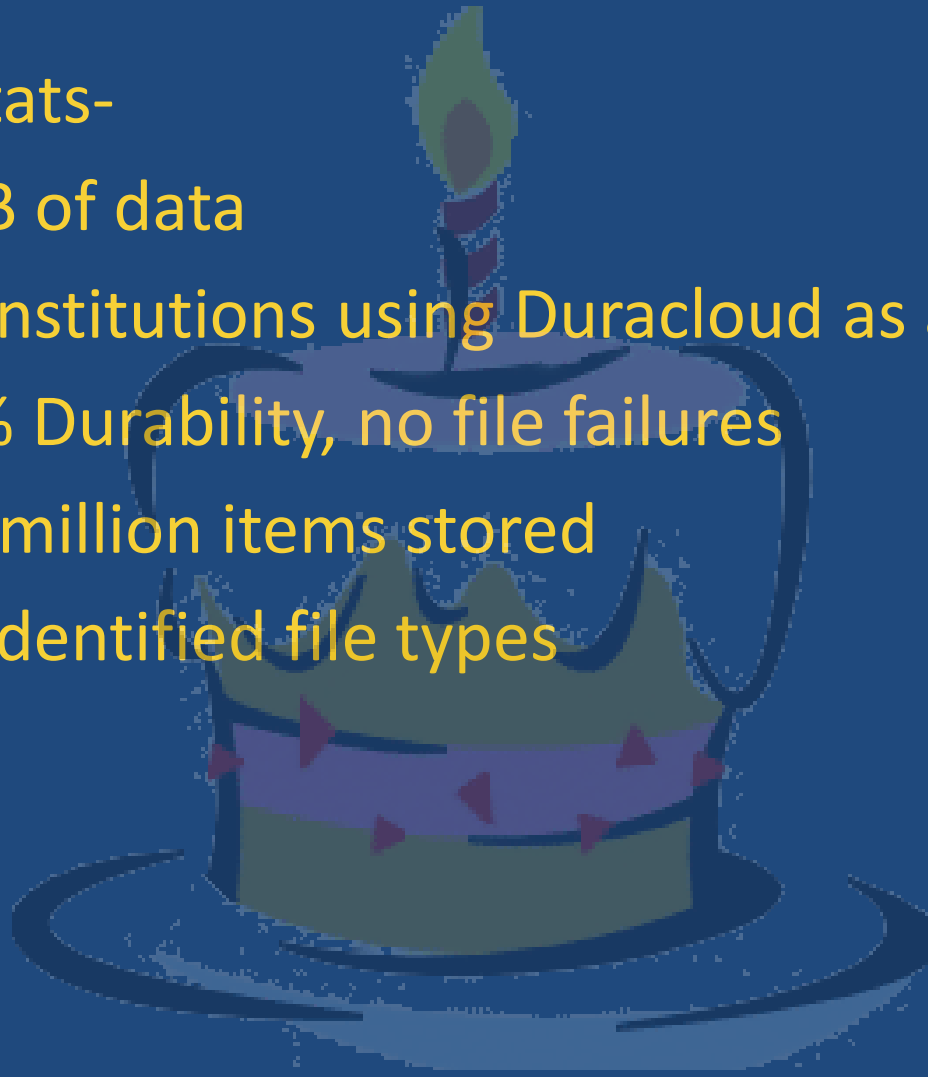
DuraCloud Roadmap

- Coming in 2013
  - ✓ Glacier integration
  - ✓ Shibboleth integration
  - Search
  - Encryption services
- Other cloud based offerings
  - Dspacedirect(hosted repository service)

Our Goal, run DuraCloud platform on academic cloud infrastructure

To find out more

- Web:
  - http://www.internet2.edu/netplus/duracloud/index.html
- Documentation:
  - http://wiki.duraspace/display/duracloud
- Webinars and Training Videos:
  - http://www.youtube.com/user/duracloudvideos
- Email:
  - csmith@duraspace.org or mkimpton@duraspace.org

DURASPACE™

DURACLOUD™