



Managing live digital content with DuraSpace services

Bill Branan
PASIG Spring 2015



DuraSpace

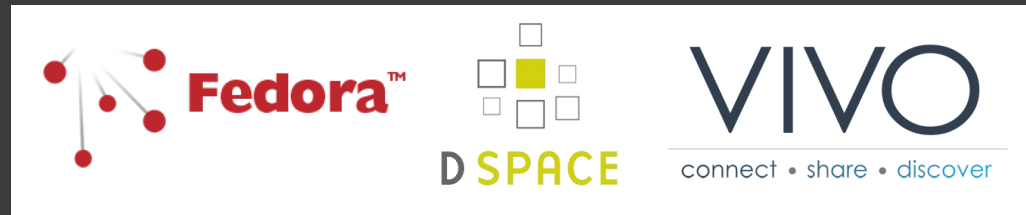
- Independent 501(c)(3) not-for-profit organization
- Mission
 - Promote durable, persistent access to digital data
 - Help to ensure that current and future generations have access to our collective digital heritage



DuraSpace

⦿ Supported technology projects:

- Fedora
- DSpace
- VIVO



⦿ Fully hosted and managed services:

- DuraCloud
- DSpaceDirect
- ArchivesDirect

DuraCloud

- Hosted digital preservation service
 - Open source code
 - Distributed offsite storage
 - Automated duplication
 - Automated synchronization tooling
 - Verified data integrity
 - Simple administration
 - Powerful integration options
 - Predictable annual billing
 - Personalized support
 - Open to anyone



DuraCloud

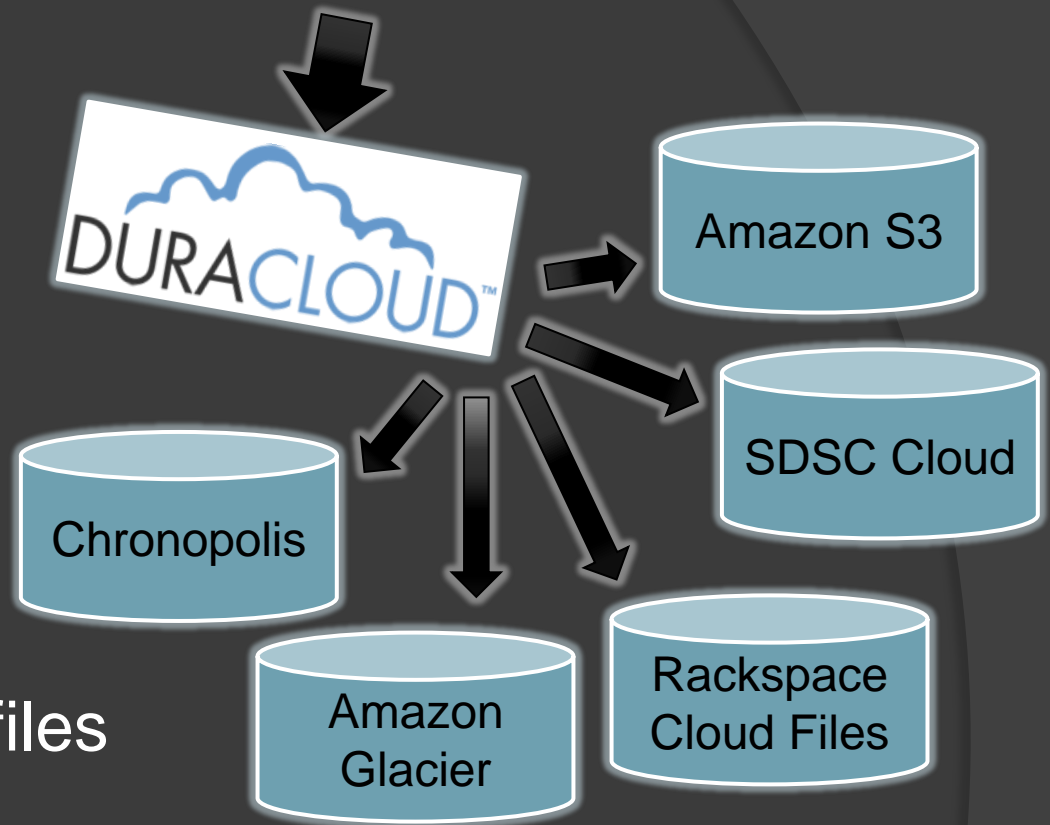
- Mediated storage

- Amazon S3
- Amazon Glacier
- SDSC Cloud
- Rackspace Cloudfiles
- Chronopolis

- Up to 15+ copies for each file stored

- Access through simple UI

- Comprehensive REST API



DuraCloud Mill

- Efficient and scalable processing of digital content
- Performs audit processing, manifest processing, duplication management, and bit integrity validation
- Automatically scales to the size of the dataset



DuraCloud in Practice

- Users span the breadth from large to small institutions
 - State libraries, state archives, institutional consortia, major research universities, liberal arts colleges, international universities, etc
- Used as primary, secondary, and tertiary storage
- DuraCloud @ TDL
- DuraCloud + Chronopolis



DSpaceDirect

DSPACE DIRECT

- DSpace
 - Open source digital repository system used by over 1500 institutions worldwide
- DSpaceDirect
 - Hosted and managed DSpace repository
 - Simple to start-up
 - Wide selection of available features
 - Free version upgrades
 - Migration support
- Automatic backup to DuraCloud



DSpaceDirect in Practice

- ◎ Supports a wide range of content categories
 - University documents
 - Government records
 - Electronic Theses and Dissertations
 - Historical photo archives
 - Audio and Video
 - Special Collections
- ◎ Used by a wide range of institutions
 - Universities, colleges, community colleges
 - American Library Association Archives
 - National Endowment for the Humanities
 - Sustainable Agriculture Education Association



DSpaceDirect Features

- ◉ Distributed deposit
- ◉ Public data access
- ◉ Managed system administration
- ◉ Unique, persistent IDs via Handle System
- ◉ Search engine optimization for Google and Google Scholar
- ◉ Permissions management
- ◉ Usage statistics and reporting
- ◉ Integration with Google Analytics
- ◉ Customizable approval workflow
- ◉ Customizable deposit license for content submissions
- ◉ Batch metadata editing
- ◉ Annual upgrades
- ◉ Personalized support



ArchivesDirect

@archivesDIRECT

- Brand New!
- Archivematica – workflow tool for generating preservation ready AIP packages
 - Customizable OAIS-based workflow processing
 - Permanent ID generation
 - Checksumming
 - Virus checking
 - File format identification and validation
 - Metadata extraction and generation
 - Normalization
- DuraCloud – preservation storage



@archivesDIRECT



DSPACE DIRECT



Chronopolis

Amazon S3

SDSC Cloud

Rackspace
Cloud Files

Amazon
Glacier

DPN



Lessons Learned

- No single service can support all use cases well
- Cloud is an excellent hosting platform for supporting preservation activities
- Preservation of digital content is becoming more accepted, and even expected

Thank You!

● Contact

- Bill Branan: bbranan@duraspace.org

● Web sites

- DuraSpace: www.duraspace.org
- DuraCloud: www.duracloud.org
- DSpaceDirect: www.dspacedirect.org
- ArchivesDirect: www.archivesdirect.org

