Integrating Fedora with DuraCloud

DuraSpace Webinar: 11 Jan 2012

Chris Wilper
DuraSpace
Who is DuraSpace?
What is DuraCloud?

- Open technology and managed service that utilizes cloud infrastructure for preservation support and access services
  - Upload once and make multiple copies to various storage providers
  - Access anywhere from any Internet connected device
DuraCloud Use Cases: A Snapshot

- Archiving and Preservation of Repository Content
  - Backup (online, off-site)
  - Restore (from cloud to local)
  - Synchronization (between local and cloud)
  - Duplication (between multiple cloud stores)
  - Integrity Checking
- Multimedia Access
  - Image Viewing
  - Media Streaming
DuraCloud Customers Using Fedora

- Colorado Alliance of Research Libraries
- Columbia University
- Discovery Garden
- Hamilton College
- Northwestern University
- University of Prince Edward Island
Typical Fedora Backups

Database & File-Level Backups
Typical Fedora Restores
CloudSync Backups

Fedora

Fedora CloudSync

DURACLoud

Object-Level Backups
## Summary

<table>
<thead>
<tr>
<th>Feature</th>
<th>Typical Solution</th>
<th>CloudSync + DuraCloud</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Backup Granularity</strong></td>
<td>All Objects</td>
<td>Any Subset</td>
</tr>
<tr>
<td><strong>Restore Granularity</strong></td>
<td>All Objects</td>
<td>Any Subset</td>
</tr>
<tr>
<td><strong>Skill Level</strong></td>
<td>Advanced</td>
<td>Basic</td>
</tr>
<tr>
<td><strong>Downtime</strong></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Offsite</strong></td>
<td>DIY</td>
<td>Yes</td>
</tr>
</tbody>
</table>
What about Speed?

- With any offsite storage, network bandwidth is a factor.
- In-house approaches can be faster for whole backups and restores (e.g. LVM or ZFS snapshots)
- But consider the entire turnaround and downtime (involving people time, support tickets, etc):
  - For in-house backups
  - For in-house restores
CloudSync Demo

- Live demo will show:
  - Quick Installation
  - Adding a User
  - Stores: Where Fedora data is kept
  - Sets: Objects of interest for a Task
  - Example Tasks:
    - Complete Backup via Copy Task
    - Selective Restore via Copy Task
  - REST API
Jan 2012: CloudSync 1.1

- Copy Externally-Managed Datastreams
- Local Directory-based Stores
- Ability to Edit Sets
- Complete REST API Documentation
Future Ideas

- True “Sync” Task
  - One-way
  - Two-way
- Batch Modify Tasks
  - XSLT
  - Java
  - JRuby/Jython
- DSpace-based Store
More About CloudSync

- **Homepage**
  
  https://wiki.duraspace.org/display/CLOUDSYNC

- **Latest Downloads and Documentation**
  
  https://wiki.duraspace.org/display/CLOUDSYNC10

- **Feature Requests & Bug Reports**
  
  https://jira.duraspace.org/browse/CLOUDSYNC

- **Mailing List**
  
  https://groups.google.com/group/fcrepo-cloudsync
More About DuraCloud

Web site:
www.duracloud.org

Email:
Carissa Smith, Partner Specialist
csmith@duraspace.org

Introduction to DuraCloud:
www.youtube.com/watch?v=rNB8l2IMQy0

Entire DuraCloud Webinar Series:
www.youtube.com/user/duracloudvideos

Email Carissa if you are interested in a free live DuraCloud demonstration.
Free DuraCloud Trial Accounts

Sign up today and get two months of DuraCloud FREE
Photo/Icon Credits

- Hard Drive: http://tango.freedesktop.org/Tango_Desktop_Project
- Folder: http://www.openclipart.org/detail/13740
- Database: http://www.openclipart.org/detail/68413
- Checkmark & Delete: http://veryicon.com/icons/system/on-stage/