2014-09-04 - PSU Beta Pilot

Time/Place

This meeting is a hybrid teleconference and IRC chat. Anyone is welcome to join...here's the info:

- Time: 11:00am Eastern Daylight Time US (UTC-4)
- U.S.A/Canada toll free: 866-740-1260, participant code: 2257295
- International toll free: http://www.readytalk.com/intl
    - Use the above link and input 2257295 and the country you are calling from to get your country's toll-free dial-in number
    - Once on the call, enter participant code 2257295
- IRC:
    - Join the #duraspace-ff chat room via Freenode Web IRC (enter a unique nick)
    - Or point your IRC client to #duraspace-ff on irc.freenode.net

Attendees

- Adam Wead
- David Wilcox ⭐
- Andrew Woods
- Unknown User (escowles@ucsd.edu)
- Michael Durbin
- Robin Dean
- Ed Fugikawa
- Carolyn Cole
- Hector Correa
- Betsy Coleman
- Chris Beer
- Eric James

Agenda

1. Introductions
2. Update on current status of PSU pilot work
3. Discussion of open issue
4. Next steps
5. ...

Minutes

Introductions

- Andrew Woods
    - DuraSpace
    - Fedora tech lead
  Esme Cowles
    - UCSD
    - Working on F4 dev for 1.5-2 years
  David Wilcox
    - DuraSpace
    - Fedora Product Manager
  Mike Durbin
    - UVa
    - Working on F4 for a year
    - In the Fedora community for many years
  Robin Dean
    - Colorado Alliance
    - Interested in migrations from F3
  Ed Fugikawa
    - Colorado Alliance
    - Tech Lead
    - Running lots of Fedoras
  Carolyn Cole
    - PSU
    - Scholar Sphere on F3, moving to F4
  Hector Correa

- PSU
- New to Fedora

Betsy Coleman
- University of New Hampshire
- Software Developer
- Grant to integrate OpenGeoportal with Fedora 4
- Existing F3 installation

Chris Beer
- Stanford
- Developer

Adam Wead
- PSU
- Developer
- Worked at RockHall, built Hydra app
- Migrated Hydra from F3 to F4

Eric James
- Yale
- Programmer/analyst
- Working on Fedora dev for the past year

PSU beta pilot

- Rolling out ScholarSphere 2.0
  - In production for about 3 years
  - Migrating to F4
  - Using Sufia

Getting Sufia working with F4

- Mostly done
- Justin Coyne moved ActiveFedora to F4
- Descriptive metadata (in XML) moved to RDF properties on nodes
- Can create collections
- File uploading
  Intermediate node creation
  - ModeShape requires intermediate nodes for performance reasons
  - Andrew Woods: ModeShape Beta 1 (in F4 B3) makes progress toward possibly negating the need for intermediate nodes
  - Adam Wead: Using NOIs as IDs. Cut into tuples, each tuple is a node

Versioning

- In F3, versions are persisted in a content data stream
  - Need to preserve these versions
- Need to restore previous versions as well
- Michael Durbin: Do you need to preserve the original dates?
  - Adam Wead: Yes
  - Michael Durbin: Need to do a code change; lastModified date will be the date you import into F4
  - Could iterate through old versions, tag each with a date
- Robin Dean: Would dates for all objects be changed on import?
  - Michael Durbin: Yes, but this will likely change as we build tools for migrations
- Chris Beer: Using a metadata field for the date might make more sense than the system modified date - that date might already be misleading
  - This would also provide more flexibility - you could have a user modified date and a system modified date for example
- Restoring versions: The system creates a new version just before you restore a previous versions
- In F3, does changing a datastream create a new version of the object?
  - LastModified date gets updated
- In F4, when you version something, everything below it in the graph also gets versioned
  - Lots of flexibility: you can version on every change, or just on particular changes

Auditing

- Sufia conducts audits of repository objects
  - Now that F4 can do this, should we take advantage of this functionality?
  - Could update batch jobs to call fixity service
    - Probably faster

Batch jobs

- Uploaded files added to batches
- Maybe use weak references?

Datastreams

- Need to represent rights metadata in RDF
- Need to decide on an ontology

Actions

- Schedule a monthly meeting
- Andrew Woods will send out a note ahead of Hydra Connect to see if there is a need for a discussion prior to that event