2015-04-23 - Fedora Tech Meeting

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

- Unknown User (escowles@ucsd.edu)
- Nick Ruest
- Andrew Woods
- A. Soroka
- Unknown User (acoburn)
- Jared Whiklo ⭐
- Stefano Cossu
- Mohamed Mohideen Abdul Rasheed
- John Doyle
- Justin Coyne
- Unknown User (daniel-dgi)

Agenda

1. Audit sprint progress
2. Blank node updates
3. F4 export/import and content-negotiation
4. LDP W3C workshop this week
5. Islandora
   a. Fedora 4 IG call, Friday, April 24
   b. Performance testing triplestores for Islandora community recommendations
   c. auditTrail mappings feedback
6. Unable to locate Jira server for this macro. It may be due to Application Link configuration. - AIC use case

Minutes

Audit spring progress

- End of second audit sprint. Focusing on internal audit implementation that will create nodes for audit events. Implemented and working, tying up some loose ends.
- Final issues: using direct access to underlying kernel, need addition integration tests, some packaging issues for the vagrant. Looking to wrap these up by the end of sprint.
- VM with that desired configuration. Unable to locate Jira server for this macro. It may be due to Application Link configuration.
- Once this is ready, prepare for a new release. Would this version be 4.2.0 or 4.1.2?
  - Java 7 to 8 will cause a release version bump as well.
  - Chance of next release might be non-backwards compatible (relating to blank node discussion?)
  - Esme Cowles: I like 4.2.0, but also not sure on when Java 8 will be ready.
  - A. Soroka: Java 8 is scheduled for beginning of May.
  - Andrew Woods: Significant changes coming, so probably be better to have a couple releases to ensure each gets reviewed/tested properly.
- Documentation is coming together, still some configuration and a top-level feature page.
- Fixity events still not at all implemented. Possibly as a separate topic once Ben Armintor produces a proposal for a re-factor of fcrepo LDP fixity service.
- Audit sprint team is complete other than responding to review comments.
- Ben appears with some ideas on new fixity implementation.
  - Fixity resource needs to be a container that you can POST to get new checks and GET past checks results.
  - Like the fcr:metadata resource. Rob (Sanderson) thinks the metadata resource could be the fixity container.

Blank node updates

- What is going on? What is in the master branch? What clarifications need to be made?
- Justin Coyne: Skoleminization of blank nodes is a feature we rely on and if true bnodes are created a lot will have to be changed in our workflow.
- Andrew Woods: Skoleminization is happening, but nodes are not published as first-class resources.
- A. Soroka: You can update properties on bnodes. Nodes are always skoleminzed, we are talking about public exposure of this skolemnization. You can still update them.
- A blank node will always be the object in a triple with a resource as the subject. To update that blank node you apply a SPARQL PATCH to the original subject node with the changes to the object (blank node).
- Justin Coyne to open a ticket to update documentation of how to update blank node properties.
- Justin Coyne: I feel that this change is large enough to not be a point release, when are we ready for 4.2.
- Andrew Woods: Depends if Audit service is ready to go. Possibly this weekend, definitely in single digit days.

F4 export/import and content-negotiation

- Import/Export on API. Currently they are scoped at the JCR level, many triples generated by Fedora layer of application, most useful stuff is at that level. If you export you get the raw data.
- Export functions are robust as you can export an entire tree with or without binaries. As the project is moving towards being more standards based and removing cruft. Should we remove existing import/export and use a GET request for export with content-negotiation and PUT/POST for import.
- This would limit the tree export, possibly one level down with a Prefer header. What would not be there is to get all the binaries associated with some resources, you would get the references to those binaries.
- The change we are making here is to support these operations via more normal HTTP requests. Keeping it more cleanly about what you get in the export.
- A comparison would be beneficial to see what the differences would be between existing functionality and this new HTTP export functionality. To what degree are people depending on existing export.
- fcrepo-camel is another option for bulk operations.

LDP W3C workshop this week

- Hydra/Fedora communities were well represented.
- LDP charter is coming due at end of July. LD Patch and LDP Paging also come due end at the of July. LDP paging is likely not moving forward, not enough adoption. LDPatch will likely show up as a recommendation.
- New LDP charter will be drafted to help shepard that project along.
- One of the goals of the group will be defining a consistent and formalized pattern for extending the capabilities of LDP.
- Other attendees of the meeting included Oracle, MIT, IBM, Greg Kellogg (Ruby), and a couple independents.

Islandora

- Fedora4 interest group meeting tomorrow, open invite to all.
- We have an issue for performance testing triplestores, still working through how to test. The community might also be interested.
- Audit trail event mappings completed and sent to the listserv for feedback.
- Do they work for everyone now? Some comments related to content removal. Also around fixity mapping (ie. should do premis:fixityCheck).
- Esme Cowles: Eric James suggesting using premis:removal to differentiate between objects removed and files removed. Which is a good idea. Currently we are using premis whereever there is a good match. However there is not an event that maps to editing metadata or removing a file, etc.
- Nick Ruest is hoping he has got all the audit trail predicates mapped, seems to be complete and no one has mentioned any missing ones.
- Danny Lamb issued a PR for migration stuff, applying RELS-INT to the non-RDF resource description.
- Audit trail mappings will have to be added to migration-utils and be tested.
- These will be the default migration paths for audit trails.

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

- Desire to have a prefer header to expose the properties of related objects to the resource being requested.
- Use case is of syncing master repository to a secondary repository with smaller subset of data.
- This takes several calls right now to get the resources and all the properties and would be better if you could get more data per request.
- Expanding on the existing GET prefer header (return=representation) to get properties on the extended result set.
- A. Soroka warning that this might be an expensive operation depending if the resources are indexed in Modeshape.
- Examples of possibly problematic data would be across federated file system or a "projected query"?
- This is akin to asking for either a bounded graph or explicitly defined list of resources.
- Stefano Cossu mentioned they are not using an external index because they want the resources in the state at-that-moment and can therefore not afford the delay of indexing.