

2013-09-26 - Fedora Committer 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 \(UTC-4\)](#)
- ReadyTalk:
 - 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

- [Greg Jansen](#)
- [Esme Cowles](#)
- [Ye Cao](#)
- [Markus Haarländer](#)
- [Matthias Walter](#)
- [Eric James](#)
- [Michael Durbin](#)
- [A. Soroka](#)

Agenda

1. 20 - Review current fcrepo4 work.
 - a. [Triple store integration](#)
 - b. [Customizable search indexes](#)
 - c. [Large file performance](#)
 - d. [Fcrepo3 migration](#)
 - e. [roles-based AuthZ](#)
2. 10 - Scalability of the API for children
 - a. Moving away from using Jena Model as container for triples in the kernel
 - b. Moving to a streaming API all the way through the kernel
3. 10 - Stakeholder needs in a customizable search indexer (Solr)
4. 15 - [Wiring and configuration](#) (see also [Comparison - Wiring](#))

Previous Actions

- [Michael Durbin](#) to create Pivotal tickets for UI polish
- [Benjamin Armintor](#) to send email for others to sign his and other committers' keys

Minutes

1. reports on current work
 - a. triple store integration - Esme
 - i. will be looking at Mulgara and JMS indexer to incorporate non standard SPARQL queries
 - ii. looking at multiple levels based on reviewing use cases
 - iii. kitchen sink now has triple store (Question for Andrew: Is vm built from master automatically)
 1. Answer from Andrew: "Yes, Jenkins deploys kitchen-sink on each commit to: <http://futures6.fcrepo.org:8080/fcrepo/rest/>"
 - b. large file performance - Eric
 - i. 3 avenues explored
 - ii. researching different configuration settings (e.g. caching hashes showed dramatic improvement)
 - iii. Scott commented on how fixity done in Fedora3
 - iv. fyi, looks like we actually read the data and do a new checksum when we do the [fcr:fixity](#) check, the low-level code that does the reading/checksumming is here:<https://github.com/futures/fcrepo4/blob/master/fcrepo-kernel/src/main/java/org/fcrepo/kernel/Utils/LogLevelCacheEntry.java#L92> (Esme)
 - c. migration - Mike
 - i. 90% completed and the remaining 10% is slow
 - ii. working on display, especially focused on performance when many children as displayed
 - d. roles based - Greg
 - i. will be working on integration test needed for effective roles for tree object
 - ii. roles can have expanded permission set (Scott)

- iii. XACML will be used for each role
- iv. policy enforcement point
- 2. API for children
 - a. effective hierarchy needed to maintain good performance
 - b. a couple of tests suggested (Greg has the action items)
- 3. search indexer - Markus
 - a. external indexer (esp. Solr)
 - b. [Design - Customizable Search Index](#)
 - c. Scott described how they use Solr in F3. change METS to Solr XSLT to customize
 - d. what should be indexed? data streams, text extraction from spreadsheet, PDF, etc.
 - e. Possibly use METS which can point to binary files
 - f. maybe use a default Solr schema; consider exposing to public and document how to use
 - g. alternatives to XSLT should be considered
 - h. need to keep the idea of distributed hardware (e.g. separate Solr and Fedora)
- 4. wiring - Adam
 - a. changes will be hard and are hard now
 - b. future proofing may not be as important as buy in from the group
 - c. examples used currently or in the past: struts, Spring, JAX, CDI
 - d. feedback from the community needed for <https://wiki.duraspace.org/display/FF/Comparison++Wiring>

Actions