2014-02-20 - 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)
- 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

- frank asseg
- Ed Fugikawa
- Andrew Woods
- Stefano Cossu
- Nikhil Trivedi
- Scott Prater
- Osman Din ⭐
- Eric James
- Greg Jansen
- Michael Durbin
- Mike Daines
- Ben Pennell
- A. Soroka

Agenda

1. 4.0 Features: review of responses

Feature	Votes (out of 7)
Authorization	5
Backup	4
Batch Operations	1
Clustering	5
Content Modeling - Structural	5
Managed External Datastreams	5
Store/Deliver Large Files	6
Search	7
Transactions	3
Triplestore	6
Versioning	4
Non-Functional: Easy Deployment	6
Non-Functional: Performance - Single-node	4
Non-Functional: Performance - Clustered	3

2. Resource locking model
3. Configuration and support of both protected and public endpoints (Greg Jansen)
4. Bring the meat ax down on our leakage of JCR abstractions through our API and RDF (Michael Durbin)
5. Fedora ontologies and concept mappings
6. other...?

Minutes

1. **4.0 Features: review of responses**

Andrew and Greg: The sent out features e-mail did not get any negative responses. The feature list covers items that were agreed upon by stakeholders and developers.

The idea was to check was to: 1. check off 2. raise any features not on the list, or re-think

The idea was to prioritize features that are most important to developers (representing institutions send out a broader message, saying that these are the features for 4.0 which drive the focus of development, in order to get sign off from stakeholders.

The page shows where the preferences lie (not to say that batch operations or transactions won't be worked on, just the priorities).

Greg: With respect to clustering, there are two distinct use cases: (1) performance; (2) redundancy. Wondering if we have emphasized that. There are some expectations in the community about redundancy.

Andrew: Yes, two distinct things. Wondering if we talked about. . .perhaps only APTrust has hinted at that at (some point). Do other folks think that's true?

Response: makes sense to talk about it

Andrew: Any other suggestions?

Esme: In our case (and as mentioned in e-mail), a lot of these features are implemented in the front-end (via hydra). So, e.g., even if F4 beta has no auth, it'd still be an acceptable product.

Andrew: These discussions are taking place . . .exploring the idea that Hydra starts to leverage F4.

Martin: Same applies to Islandora.

Andrew: I encourage conversations with Islandora as well. With David Wicox on board, we should get these conversations going. There are different perspectives when it comes to the repo. One group wants to push functionality down to the lower level, and other group wants a minimal repo that focuses on key cases, such as preservation. It is up to us to draw the right line.

Esme or Scott: Auth could be externalized. Features such as versioning, search, content modeling are good.

Andrew: Let's continue the discussions. After discussion with stakeholders, we have this list.

Stefano: People might find that F4 is a completely different product from F3 in terms of features. Lack of features may put implementation burden if some features are absent (e.g. does auth have to be implemented in client?).

Andrew: Agreed. But the first focus of 4.0 release beta is not upgrade from Fedora 3.. subsequent releases, 4.1, 4.2, will make it easier to upgrade.

2. **Resource locking**

Mike summarizes the problem. There's some locking which is not there and can be misleading if it's assumed to exist. We need to explore JCR locking functionality.

Adam: discusses 2 different proposals about how to handle these locks.

Mike: No personal preference, but need to do something that works

Adam: Resources that exist in repo, or in http api . . .

Andrew asks whether JCR spec talks about node being creating representing a lock.

Adam: No, it's an implementation concern.

Andrew: So, it's a question of where does this thing reside

Mike: If you delete the node, the lock disappears or not?. . .

Adam: Not sure...

Mike: We can figure out as we implement this

Andrew: Any design ideas before you code that you can share would be good on IRC and pivotal ticket . . .

3. **Configuration and support of both protected and public endpoints (Greg Jansen)**

Andrew summarizes Greg's point about two war files: (1) standard file with no auth enabled by default; (2) war file with a repository.json to point to auth classes and web.xml specifying protected resources.

Main questions: (1) how do we support auth out of the box; (2) does it make sense to push these challenges to the application layer versus the servlet container.

Esme discusses details about an auth implementation for a file delivery application at UCSD (contained a parameter to present a login challenge). Says having a single policy would not allow people to do what they want to do with auth.

Andrew: Having 2 end points is not sufficient. E.g. in a the case of some hierarchy, where the parent is not open to public but the images should be.

Greg: Haven't thought about suppressing the challenge. Having 2 end points doesn't seem RESTful.

Esme: Having a 401 and 403. 403 was the default in the file delivery auth layer.

Greg: Makes sense; the session could be checked.

Andrew: Doing this in application, not in servlet container. Create a ticket to explore.

Greg: sure.

4.Bring the meat ax down on our leakage of JCR abstractions through our API and RDF (Michael Durbin)

Andrew: There's some discussion out about exposing JCR functionality such as providing JCR search capability. The general notion of abstracting away jcr is good. For example, if we throw away ModeShape. . .

Mike Durbin: should be prioritized

Andrew: Makes sense; create ticket; review.

5. Fedora ontologies and concept mappings

Andrew: It's ongoing work that primarily Adam has been engaged. Lots of different aspects to it, and we won't go into them . There are a few different notions on the table floating around and often times intermingled. We have talked about trying to map fedora concepts to JCR concepts and

presumably in a higher in the chain, specifically LDP. How doe we expose fedora concepts in LDP and provide a mapping. Adam has been working on creating a model of these ontologies. The goal is to speak to the community clearly.

Asks Adam whether he has anything to add . . .

Adam: No, that was fantastic.

Addendum: The mapping from Fedora to LDP is available in OWL here: https://github.com/futures/ontology.

Andrew: Thanks . . . Reminder: last day of sprint . . . any objections if we do the sprint wrap up at 4 pm est? See you then.

New Actions

- Greg Jansen to create ticket for exploring "parameter-based" AuthZ
- Michael Durbin to create ticket for cleaning JCR references out of F4 responses