

2016-09-08 - 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)
- Dial-in Number: (712) 775-7035
 - Participant Code: 479307#
 - International numbers: [Conference Call Information](#)
 - Web Access: <https://www.freeconferencecallhd.com/wp-content/themes/responsive/flashphone/flash-phone.php>
- IRC:
 - [Join the #fcrepo chat room via Freenode Web IRC](#) (enter a unique nick)
 - Or point your IRC client to #fcrepo on irc.freenode.net

Attendees

- [Longshou Situ](#)
- [Esmé Cowles](#)
- [Yinlin Chen](#)
- [Bethany Seeger](#)
- [Aaron Birkland](#)
- [Jared Whiklo](#)
- [Jim Coble](#)
- [Andrew Woods](#)
- [Jennifer Lindner](#) ★
- [Michael Durbin](#)
- [Doron Shalvi](#)
- [A. Soroka](#)
- [Daniel Lamb](#)

Agenda

1. Updates and Announcements
 - a. [Archivematica Camp](#) - updates? outcomes?
 - b. Import/Export [sprint update](#)
 - c. [Bill Branan](#) and [Danny Bernstein](#) will likely be giving the "many members" issue a look
 - i. <https://github.com/fedora4/fedora4/commits?author=quarles>
 - ii.
 - iii.
 - iv. Unable to locate Jira server for this macro
2. Application Link configuration. [son's comment](#)
Unable to locate Jira server for this macro. It may be due to
3. [Fedora 4.7.0 planning](#)
 - a. Depends on
4. Import/Export: Clarify
5. gradle in fcrepo-exts?
Unable to locate Jira server for this macro. It may be due to Application Link configuration.
6. ...
7. Status of "in-flight" tickets

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due
Unable to locate Jira server for this macro. It may be due to Application Link configuration.										

Ticket Summaries

1. Please squash a bug!

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due
Unable to locate Jira server for this macro. It may be due to Application Link configuration.										

2. Tickets resolved this week:

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due
Unable to locate Jira server for this macro. It may be due to Application Link configuration.										

3. Tickets created this week:

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due
Unable to locate Jira server for this macro. It may be due to Application Link configuration.										

Minutes

Agenda

1. Updates and Announcements

a. Archivemata camp

The first question was, is there any actual work that might be done with Fedora? And the answer is integrating at storage level is not good idea. Archivemata does lots of copying, moving bits, more like file system and not like a repository -- Fedora won't work well there.

However, how do you move stuff from Archivemata to Fedora? Let's let Nick and crew get sprints done and we'll have most of that done -- Archivemata produces bags and Fedora is working on accepting them.

The last is with the format policy registry. Lots of folks are interested in breaking it out of Archivemata, and amongst the people interested are [JISC](#), in the UK. They are preparing to manage an Archivemata service. So, let them get done before discussing breaking it out. (Context: many people who are using Fedora like Archivemata's format registry but may or may not want Archivemata itself, including Islandora -- the community would have interest). [Justin Simpson](#) promised to be in touch and restart conversation when they are ready.

In sum though, the vast majority of discussion was inward facing and around community development.

b. Update from import/export:

[Esmé Cowles](#): We've made good progress on export, and got basic import utility merged in, and are working on round trip for phase one objectives, w/o doing anything with bags. We're working on sample data sets and driving initial docs, on how different this utility is from RDF serializer Camel project that was the starting place for project.

We're in a good place for future sprints to tackle bugs and other features. We'll have an executable jar that addresses objectives and docs for how to use utility, enough so people will be able to at least give it a try and provide feedback that could get rolled in to later sprints.

c. Two Java devs, [Bill Branam](#) and [Danny Bernstein](#), will have time to address Fedora performance issues around having many members. There's a Hydra change of direction (from members to collection) PR in, which is a mitigating change, but this will directly address this issue. (Status of the PR is the commits are not in but supporting commits are, Curation Concerns PRs are priority right now. There were some concerns, around whether solr indexing would change, but (AFAIK) concerns were addressed in it. PR should merged in next few weeks.

2. Revisit default behavior on PUT requests containing RDF bodies

The issue is a request for changing the interaction contract on PUT requests with Fedora. What we're asking for is changing the model from how it relates to server managed triples. In the current approach, if you want to do a PUT, replacing all triples, you can make a PUT with a complete representation of the resource, with both user and server managed triples in it. If you don't have server managed triples you get a 409 in response. The change being requested is for clients to only include user managed triples and if they include the server managed ones they get the 409.

[A. Soroka](#): For context -- what [Stefano Cossu](#)'s trying to do calls into question whether or not he should be using PUT at all, and he shouldn't be -- for these reasons. Here's his use case. Stefano doesn't care what the current state of resource is, he wants to mutate it (he may or may not want to remove properties) with no questions asked. He's only interested in the triples he actually wants to change. Here's why it matters: he's arriving with a request in hand, fully formed (adds or removals), -- and it's not a complete representation. It's a delta, or a future state of resource, however we want to describe it, and this is a direct conflict with what's a PUT. He's doing a PATCH. He's not doing it through a SPARQL update because it's unpleasant -- what he wants is a more convenient way to PATCH, like with RDF Patch or LDP Patch. Whether or not we act on this is up for discussion, but I feel we have to consider whether to do this with PUT.

[Andrew Woods](#): We need to verify that's correct description of Stefano's use case. And we need to reaffirm we're doing what we need to with PUT.

[Aaron Birkland](#) -- The only time I've come across Stefano's use case is in integration tests, where I don't care what the state of resource is. And that's kind of minor, so I understand the appeal of PUT for this kind of situation. That said, I think Fedora's current behavior is correct. The LDP spec is explicit on what clients should do with triples they don't care about (PATCH them through), but I can see desire for shortcuts in certain situations.

Adam: Stefano does care about the resource in the sense of he wants it to be maintained except for his changes, he just doesn't care what that state is. I respect his claim that it could be easier.

Andrew: We could just add a header -- is there another way to allow GET/PUT to work? -- It would require the client requesting minimal state -- we'd have to change the current get to support change in PUT.

Doing PUT w/o server managed triples, a user can do that with an additional header: `Prefer: handling=lenient; received="minimal"`. Both cases are handled in PUTs. You can pass in headers to get what you want. So, we'd just be losing one scenario possibly.

Does anyone know if there's a way to allow server managed triples?

There are references to it in two places in the LDP spec, our current interaction supports one scenario. And Stefano's ticket is supported in a may clause in the other.

Andrew: HTTP PUT should be idempotent and currently is not, because server managed properties change. (Later IRC discussion indicated that in fact current PUT behavior is at least arguably idempotent for a definition of that term derived by [Jared Whiklo](#) from the HTTP spec.)

Adam: How would we get around it? Stefano's request would.

Andrew: The spec says we should be able to make three identical PUTs and all three succeed and the state is the same regardless if you did it once or three times.

Adam, Andrew: The state of what, though? The server managed triples? Date modified would change .. but the state of resource is not the state of representation, so if change appears in representation but not resource is this a problem? HTTP is concerned only with representations. The representation has not changed, is idempotent but the resource has changed, date has, etc. Also concerned with does it succeed or not.

Aaron: In the example Adam gave before, with three same PUTs in minimal mode, you're modifying full representation three times. That's not what you'd expect with idempotent behavior, but comparing the two different representations of the resource. But where the current behavior is now (doesn't matter if they all succeed to meet criteria of idempotent), is clearly idempotent.

(a little mind bending)

Aaron: Idempotent is state, responses are not relevant here.

General agreement - we should come back to Stefano with we think Fedora's current implementation is correct, however, what are things you'd like to see to meet your needs? (Maybe an easier way to PATCH, etc.)

3. About Fedora 4.7 release

Andrew: I have two questions. One, is there any reason that we'd want to hold off on next release? The main thing it does is move Fedora to latest version of ModeShape, ModeShape 5 (which will help with Infinispan). And two, which is a blocker -- making sure mechanisms for migrating existing data work. Are there any other reasons we should hold off on 4.7?

[No response.] Seems to be no.

(general discussion) In fact we should release soon and let people who want to use ModeShape 5 use it more easily and will lead to better testing (results will be limited to mode 5 bugs).

Esme: I agree it's a blocker but we need someone with real production data to do it and find bugs.

[Michael Durbin](#): We won't find problems before, no one will do it. We'll have to ship and let people find problems and we'll fix them.

Andrew: Would it be useful to have a small utility for people to do migration?

[Bethany Seeger](#): Not sure if a script would be useful, but maybe just flesh out the fcr:backup/restore page a bit more?

Andrew: Considering a RC candidate would require a release manager -- who can do it? Adam?

Adam: I did 4.6 so I could avoid 4.7.

Andrew: We'll think about who can do it, and talk about how to make it happen.

4. import/export - Clarify?

Esme: No, I understand, no need to discuss.

5. Gradle in fcrepo-exts?

Bethany: Aaron Coburn and I've been talking about the possibility of [Gradle](#) in some projects in fcrepo-exts, instead of Maven. It does same thing as Maven, pckgs things. We want to see what the community thinks about some projects in fcrepo-exts using gradle.

general: Gradle uses Groovy language, Jared notes that [Unknown User \(acoburn\)](#) has seen that pax exam tests can be done but are more difficult.

Andrew: First we need to make sure Gradle does all that Maven does and that it won't impact fcrepo in some unknown way, by having some projects in fcrepo-exts using it.

Adam: Let's continue Gradle conversation next week, tell Aaron it received a thunderous 'maybe'.