

2017-01-12 - 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

- [Danny Bernstein](#)
- [Esmé Cowles](#)
- [Bethany Seeger](#)
- [Kevin Ford](#) ★
- [Jim Coble](#)
- [Andrew Woods](#)
- [Yinlin Chen](#)
- [A. Soroka](#)
- [Benjamin Armintor](#)
- [Jared Whiklo](#)
- [Daniel Lamb](#)
- [Ayse Durmaz](#)
- [Doron Shalvi](#)

Agenda

1. [Fedora 4.7.1 release status](#)
 - a. Need database testing
 - b. Need backup/restore testing
 - c. Targeting release date of: Jan 23
 - d. Minutes
 - i. Bethany update: Database testing is required. Confirmed targeted release date of 23 January.
 - ii. Jim Coble at Duke: Currently running ingest script of RDF and binary resources to fcrepo4.7.1-rc2 on tomcat7 and mysql. Will add to 4.7.1 testing page.
 - iii. Andrew: backup/restore testing to-date sufficient.
 - iv. NLM (Doran?): Can assist with additional testing of RC2 in tomcat7/mysql.
 - v. Esmé: Re-running ActiveFedora tests, but issue seems to have been resolved.
2. Stakeholders for release testing
 - a. Widespread community release testing is important. To support this goal, I ([David Wilcox](#)) recommend we include a section on the release testing page that lists stakeholders using their confluence usernames, along with their institutions. These stakeholders would be representatives of institutions that are committed to testing release candidates as they become available. This would have the benefit of sending out direct notifications when a release candidate is ready for testing, and help ensure each release is tested by as many people /institutions as possible.
 - b. Minutes:
 - i. Andrew: Would like to have a list of individuals/institutions who would be willing to perform release testing as needed. They would receive notification when it comes time to test releases. Can we please have your information so we can contact you at the appropriate time?
 - ii. Esmé: Would like to see more people involved, but believe the last few releases - the testing associated with a release - have been better. Perhaps this isn't as much an issue presently.
 - iii. Andrew: Want to make sure that the tests meet the community needs, especially at scale. This is about surfacing this important need and perhaps formalizing it. We can table this for now with the caveat that a more formal process may be required in the future as needed.
3. Community planning
 - a. Pain points
 - i. [Performance of "many members"](#)
 - ii. Ugliness of pairtree resources
 - iii. ...
 - iv. Minutes:
 1. Andrew: The above two are the most prominent pain points we hear about. Interested in knowing what other significant pain points users have experienced so we can add them to this list and track them.
 2. Ben: The migration of PIDs from fcrepo3 to fcrepo4.
 3. Doran(?): Migration from fcrepo3 to fcrepo4 in general. In our case, for example, we are in the millions of resources. We want to migrate successfully and in a reasonable amount of time. We've had problems migrating a large set of objects successfully and in a reasonable amount of time. After 100K, or so, it would fail. We've been using

- migration tools, but have experienced a number of memory issues, in the migration tool and fcrepo4. (The memory issue in fcrepo4 may have been addressed since last attempt.) Can try to re-run.
4. Esmé: Loss of system dates. That is, the migration of creation/modification times from fcrepo3 to fcrepo4.
 5. Kevin: Inter-repository referential integrity can be a pain point, but it has benefits too.
 6. Adam: Inter-repository referential integrity has been a major pain point for Dan (?). Now, or at least for a time, it made loading data impossible.
 7. Danny (?): "performance of many members" issue.
 - a. Andrew: Aaron has been looking into a layer of abstraction to handle this, in part by having an identifier converter/handler between fedora and modeshape, but he has hit a blocker such that the effort has escalated considerably, if possible at all. Basically, the idea is to save enough information that the resource identified by a particular URI does not need to be dereferenced, which is the root cause of this pain point. Memento API may render this issue moot.
 - b. Danny: I have an interest in seeing this addressed, so allow me to solicit 'best steps' and I'm happy to contribute to this.
 - c. Ben: Mode functionality is such that paths are not unique. Therefore you must dereference a resource to get its path. There is some benefit to a URI translation refactor, but to fully realize the benefit you'd need to implement an alternative index for Mode. Probably something on the order of 3 weeks of work. It would have to be funded.
 - d. Andrew: Perhaps Ben and Danny could hammer some ideas out.
- b. Features of interest
- i. RDF1.1 serialization - breaking change
 1. Andrew: Came up during 4.7.1 release testing. Essentially, some datatypes were removed from the output. A patch was injected to route around this problem (i.e. return to 4.7.0 behavior) but how long to keep it around? (We want to take it out.) The question is "when"? Which release? Can the community tolerate a release that would be considered 'breaking.' Background note: we've assumed that the switch to semantic versioning will happen when the fedora software aligns with the API.
 - a. Esmé: Perhaps we switch to semantic versioning and, as the numbers climb, so be it.
 - b. Andrew: Ideally, increasing the MAJOR version number would be accompanied by very meaningful changes, especially to go to 5.0. Also, the community has expressed an interest in only one MAJOR release annually.
 - c. Kevin: I thought that those who responded on the email thread were more wiggly on how many MAJOR releases per year, leaving the door open for a second MAJOR release annually. Perhaps we consider two MAJOR releases in 2017. If those releases are scoped out now, the community benefits from the long timeline and will likely respond favorably.
 - d. Adam: The speed of development has proven challenging. We're developing quickly and making substantial changes based on community needs, but the community has also wanted fewer and more predictable changes.
 - ii. [AppleTrees](#) - only for new installations
 1. Andrew: Aaron Coburn has presented this possibility. It is designed to make sure the messiness of Modeshape pairtrees is contained to within the modeshape layer. In this way, the pairtree identifiers don't leak out of the repository. For example: if you have a fcrepo3 PID, you can create a resource with that PID, but, under the hood, an md5-based identifier (complete with pairtree path) is created and used as the Modeshape identifier. When someone requests the fcrepo3, the md5 is regenerated, resulting in the path to the modeshape resource, and the resource fetched. I've not tested it much, but it is promising. It is important to note that this is something for new installations, but does little for existing fcrepo4 repositories.
 - iii. Alignment with API spec - breaking change
 - iv. Bug fixes
- c. Mapping features to releases
- i. 4.7.x and 4.8.x
- d. Trajectory of development
- i. Community/ModeShape impl addresses bugs, aligns with API spec
 - ii. Alt-impls emerging in alignment with API spec
 - [Derby](#)
 - [Cavendish](#)
 - [Trellis](#)
 - iii. Common API and import/export facilitates transfer between impls
 - Transfer impact mitigated by "external-content"?
- e. QUICK MEETING WRAPUP (because of time):
- i. Andrew: All in all, I'd like us to consider (new) features, breaking changes, the mapping of those features/change to future releases, all the while bearing in mind the pain points. Also, we should consider related work, be it alternate implementations or add-on components.
 - ii. Meeting time elapsed, and ended, after discussing point 3.b.ii.
4. ...
5. Status of "in-flight" tickets

key	summary	type	created	updated	due	assignee	reporter	priority	status	resolution
-----	---------	------	---------	---------	-----	----------	----------	----------	--------	------------

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

Ticket Summaries

1. Please squash a bug!

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

2. Tickets resolved this week:

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

3. Tickets created this week:

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

Minutes