# 2016-11-03 - 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

- A. Soroka
- Jared Whiklo
- Danny Bernstein
- Esmé Cowles ⭐
- Marcus Barnes
- Michael Durbin
- Bethany Seeger
- David Wilcox
- Unknown User (acoburn)
- Joe Harrington
- Andrew Woods
- Longshou Situ
- Benjamin Armintor
- Jim Coble
- Yinlin Chen
- Daniel Lamb
- Kevin Ford
- Aaron Birkland
- Andy Wagner

## Agenda

1. 4.7.0 release update
    a. Related tickets:
        i. Memory constraint on /fcr:backup https://issues.jboss.org/browse/MODE-2587
        ii. Duplicate ids from /fcr:backup https://issues.jboss.org/browse/MODE-2611
2. 4.6.1 release - patch for concurrent resource creation
3. Refactoring internal interfaces: IdentifierConverter and "getTriples"
4. Code4Lib proposals
5. Major remaining spec issues: Create-on-PUT, Headers vs. URIs for atomic ops
6. ...
7. 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 |
|-----|---------|------|---------|---------|-----|----------|----------|----------|--------|------------|

2. Tickets resolved this week:

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

3. Tickets created this week:

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

# Minutes

1. 4.7.0 release update
   a. Andrew: Releases have been made, but not publicized yet.  But there may be issues with 4.7 when deleting PairTree nodes.
      i. Kevin: I don't think it's a problem with the deletes — I've been able to trigger the error with a GET
         1. May be specific to environment (Postgresql, Tomcat)
      ii. Esme: I did my backup/restore testing with Postgres and Tomcat, and I can re-test and check if I see the same errors after the data is restored
   b. Memory issues with performing backups: see https://issues.jboss.org/browse/MODE-2587
      i. We have submitted a patch to Modeshape, they will do a release, and we can then make a new 4.6.1 release to use that
   c. Duplicate IDs: see https://issues.jboss.org/browse/MODE-2611
      i. We have submitted a patch to Modeshape, they will do a release, and we can then make a new 4.6.1 release to use that
   d. Concurrent creation issue/patch
      i. There is a patch to address this, included in the 4.7.0 release
      ii. Will backport to include in 4.6.1 release along with items mentioned above
      iii. Kevin: Have loaded data into a 4.7 repository, and haven't seen the issue that I was seeing before
2. 4.6.1 patch release
   a. If the Modeshape backports are forthcoming, then we should wait for them
   b. Good to get this release out quickly so people can upgrade to it before migrating to 4.7, to avoid the backup/restore issues
   c. Should know very soon what Modeshape's plans are and be able to plan for coordinating the releases
3. Refactoring internal interfaces: IdentifierConverter and "getTriples"
   a. Danny Bernstein is pushing forward from the work Ben Armintor did recently to address performance of resources with many links to other repository resources
   b. Internal refactoring would help make that work easier, so the performance work is waiting for the refactoring
   c. Converting identifier forms is complicated and messy in the current implementation, and there's some interest in replacing the existing Guava implementation with built-in Java 8 equivalents
      i. Bigger refactoring options include removing the getTriples function, or doing a conversion strictly from one identifier form to another (without loading JCR nodes, etc.)
      ii. Doing identifier-to-identifier translation would then involve an extra step to load resources and their properties
         1. Esme: separating those two concerns sounds like a good idea to me, and could avoid performance side-effects of current process that loads resources whether you want them or not
      iii. Loading resources and generating RDF are separate issues: you can load resources if you want (and identify them with an identifier), and the implementation can generate RDF as a stream, which may or may not require loading the resources your resource links to.  Doing the identifier-to-identifier translation makes this possible.
      iv. Should we keep using Modeshape REFERENCE properties — or switch to using URIs or something else we could use to generate references without having to load the resource from the repository?
         1. We can currently get inbound references using those REFERENCE property links.  Do we lose that, or how do we reproduce it?
            a. We can reimplement it if we want, but some users don't want it or the costs
            b. Could index inside Modeshape and be able to get inbound references if we wanted to
   d. The consensus is that the identifier-to-identifier translation refactor is a good idea, and we can wait for more information before deciding on the other issues
4. Code4Lib
   a. People planning to attend:
      i. Esme

        ii. Nick
        iii. Bethany
        iv. Andrew
    b. Code4Lib proposals are due shortly
        i. Esmé: I've been talking with a few people about a data-modeling workshop, which would probably be of interest to people using Fedora
        ii. Andrew: Maybe an import/export workshop would be a good idea?

5. API Specification
    a. There has been a lot of work to move the API Spec forward, there are a couple of outstanding issues:
        i. CRUD: Do we need to specify creating resources with PUT? See https://github.com/fcrepo4-labs/derby/issues/11 and Fedora Specification
            • Benjamin Armintor has some qualms and will write them up
        ii. Atomic Batch Operations: Headers vs. URIs
    b. Talk to A. Soroka if you have questions