DevMtg 2019-05-08

Developers Meeting on Weds, May 8, 2019

Today's Meeting Times

- DSpace Developers Meeting / Backlog Hour: 20:00 UTC in #duraspace IRC or #dev-mtg Slack channel (these two channels sync all conversations)

Agenda

Quick Reminders

Friendly reminders of upcoming meetings, discussions etc

- DSpace 7 Working Group (2016-2023): Next meeting is Thurs, May 9 at 15:00 UTC. Agenda: 2019-04-25 DSpace 7 Working Group Meeting
- DSpace 7 Entities Working Group (2018-19): Next meeting is TBD
    - Last meeting notes at 2019-02-05 DSpace 7 Entities WG Meeting
- DSpace Developer Show and Tell Meetings: On hold until interesting topics arise.
- 2019 DSpace North American User Group Meeting (Sept 23-24) Call for proposals out!

Discussion Topics

If you have a topic you'd like to have added to the agenda, please just add it.

1. (*Ongoing Topic*) DSpace 7 Status Updates for this week (from DSpace 7 Working Group (2016-2023))
2. *(Ongoing Topic)* DSpace 6.x Status Updates for this week
    a. 6.4 will surely happen at some point, but no definitive plan or schedule at this time. Please continue to help move forward / merge PRs into the dspace-6.x branch, and we can continue to monitor when a 6.4 release makes sense.
3. Upgrading Solr Server for DSpace (
    a. **Auto-reindexing in Solr**: Unable to locate Jira server for this macro. It may be due to Application Link configuration.
        i. Should this only h need to reindex?
    b. e `authority` core. Unable to locate Jira server for this macro. It may be due to Application Link configuration. Or should we use s
4. DSpace I p (Tim Donohue)
    a. PR: https://github.com/DSpace/DSpace/pull/2265 (PR is finalized & ready for review)
    b. A follow-up PR will rename the "dspace-spring-rest" module to "dspace-server", and update all URL configurations (e.g. "dspace.server.url" will replace "dspace.url", "dspace.restUrl", "dspace.baseUrl", etc)
5. DSpace Docker and Cloud Deployment Goals (old) (Terrence W Brady)
    a. Update sequences on initialization
        i. https://github.com/DSpace/DSpace/pull/2362 - update sequences port
        ii. https://github.com/DSpace/DSpace/pull/2361 - update sequences port
6. Tickets, Pull Requests or Email threads/discussions requiring more attention? (*Please feel free to add any you wish to discuss under this topic*)
    a. Quick Win PRs: https://github.com/DSpace/DSpace/pulls?q=is%3Aopen+review%3Aapproved+label%3A%22quick+win%22

Tabled Topics

These topics are ones we've touched on in the past and likely need to revisit (with other interested parties). If a topic below is of interest to you, say something and we'll promote it to an agenda topic!

1. Brainstorms / ideas
    a. (*On Hold, pending Steering/Leadership approval*) Follow-up on "DSpace Top GitHub Contributors" site (Tim Donohue): https://tdonohue.github.io/top-contributors/
    b. Bulk Operations Support Enhancements (from Mark H. Wood)
    c. Curation System Needs (from Terrence W Brady)
2. Management of database connections for DSpace going forward (7.0 and beyond). What behavior is ideal? Also see notes at DSpace Database Access
    a. In DSpace 5, each "Context" established a new DB connection. Context then committed or aborted the connection after it was done (based on results of that request). Context could also be shared between methods if a single transaction needed to perform actions across multiple methods.
    b. In DSpace 6, Hibernate manages the DB connection pool. Each **thread** grabs a Connection from the pool. This means two Context objects could use the same Connection (if they are in the same thread). In other words, code can no longer assume each `new Context()` is treated as a new database transaction.
        i. Should we be making use of `SessionFactory.openSession()` for READ-ONLY Contexts (or any change of Context state) to ensure we are creating a new Connection (and not simply modifying the state of an existing one)? Currently we always use S

`essionFactory.getCurrentSession()` in HibernateDBConnection, which doesn't guarantee a new connection: https://github.com/DSpace/DSpace/blob/dspace-6_x/dspace-api/src/main/java/org/dspace/core/HibernateDBConnection.java

c. Bulk operations, such as loading batches of items or doing mass updates, have another issue: transaction size and lifetime. Operating on 1 000 000 items in a single transaction can cause enormous cache bloat, or even exhaust the heap.
   i. Bulk loading should be broken down by committing a modestly-sized batch and opening a new transaction at frequent intervals. (A consequence of this design is that the operation must leave enough information to restart it without re-adding work already committed, should the operation fail or be prematurely terminated by the user. The SAF importer is a good example.)
   ii. Mass updates need two different transaction lifetimes: a query which generates the list of objects on which to operate, which lasts throughout the update; and the update queries, which should be committed frequently as above. This requires *two* transactions, so that the updates can be committed without ending the long-running query that tells us what to update.

Ticket Summaries

1. Help us test / code review! These are tickets needing code review/testing and flagged for a future release (ordered by release & priority)

key	summary	type	created	updated	assignee	reporter	priority	status	fixversions

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

2. Newly created tickets this week:

key	summary	type	created	assignee	reporter	priority	status

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

3. Old, unresolved tickets with activity this week:

key	summary	type	created	updated	assignee	reporter	priority	status

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

4. Tickets resolved this week:

key	summary	type	created	assignee	reporter	priority	status	resolution

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

5. Tickets requiring review. This is the JIRA Backlog of "Received" tickets:

key	summary	type	created	updated	assignee	reporter	priority

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

Meeting Notes

Meeting Transcript

Log from #dev-mtg Slack (All times are CDT)

```
Tim Donohue [3:02 PM]
@here: it's DSpace DevMtg time!  As noted in #dev , our agenda for today is very light again
https://wiki.duraspace.org/display/DSPACE/DevMtg+2019-05-08
But, let's do a quick roll call, and we can see where this agenda takes us today.  If we don't have enough to
discuss we can always wrap up early

James Creel [3:02 PM]
Hi there

DSpaceSlackBot (IRC) APP [3:02 PM]
*dyelar* has left the IRC channel

Tim Donohue [3:04 PM]
Anyone else available today @here?  If not, we'll have a very short meeting :wink:
Ok, well, I'll give a brief update on DSpace 7 anyways (for the logs here).

Terry Brady [3:05 PM]
My meeting is wrapping up
I am present now

Tim Donohue [3:06 PM]
The update on DSpace 7 is that we are marching towards a "Preview Release" (think pre-Beta, it's not going to
be all the features, but it's enough to play around with and get familiar with the new UI, new features, etc)
The goals for the Preview Release is ASAP... in tomorrow's DSpace 7 meeting we'll get a sense of whether that
is this week or next week. But, I expect it very soon.
After that, we have a second Preview release lined up for just before OR2019, concentrating on polishing up
features/concepts that we'll do training on (e.g. branding process, install process, upgrade process)
The DSpace 7 team will also be using a new Estimation Strategy for the Beta release, with estimations from
multiple developers & some built in "uncertainty factors" (to buffer those estimates).  More on that coming in
the DSpace 7 meetings though...and that strategy will help us better nail down the Beta timelines
That's the basic updates that I have at this time, but I welcome questions/comments.  Obviously the DSpace 7
meeting & notes from that meeting have all the nitty gritty details.  Tomorrow's meeting page is at
https://wiki.duraspace.org/display/DSPACE/2019-05-09+DSpace+7+Working+Group+Meeting

Terry Brady [3:10 PM]
The preview release will be a static image for dspace and dspace-angular.  I imagine that we would invoke it as
follows: `docker-compose -f docker-compose.yml -f d7.override.yml -f rel.preview1.yml`
or `docker-compose -f docker-compose.yml -f d7.preview1.yml`

Tim Donohue [3:11 PM]
Good point, I forgot to mention we'll have the Preview release available via Docker.  You can also download the
tags & install manually, but the Docker image will have test content that will auto-install when you run
`docker-compose`
Ok, that's it for DSpace 7 updates
On the DSpace 6.x side, we don't have any updates as no one is actively leading that effort (it's looking less
and less likely anyone will pick this up prior to OR2019, which is only a month away)
So, moving along to other topics
@mwood: Not sure if you are around and/or wanting to give any updates on Upgrading Solr?  https://wiki.
duraspace.org/display/DSPACE/Upgrading+Solr+Server+for+DSpace

Terry Brady [3:15 PM]
For the agenda, we should probably add a reminder for  https://wiki.duraspace.org/display/DSPACE
/2019+DSpace+North+American+User+Group+Meeting
```

(@bill fyi)

Tim Donohue [3:16 PM]
@terrywbrady: yes, good point. The call for proposals is out for the North American User Group Mtg from Sept 23-24
(see that wiki link for the call for proposals along with details on dates, accommodations, etc)
Moving along to the next topic... I don't have any updates on the One Webapp PR: https://github.com/DSpace/DSpace/pull/2265 I'd still like to see this merged before OR2019, but I've been waiting on the finishing of the Entities effort & Submission/Workflow effort (both due before 7.0 Preview)
I will note that @terrywbrady and I found some oddities in Docker compose around the One Webapp PR. We haven't nailed down if it's specific to the PR though. For him, RDF doesn't work in Docker, but it works fine for me. We haven't been able to nail down the differences in our environments either yet

Terry Brady [3:20 PM]
I have no issue on MacOS or Linux. We have been unable to find what is different in our Win10 environmments.

Tim Donohue [3:21 PM]
Good point, yes it seems specific to Windows 10. Works on my Windows 10 machine, but not Terry's
In any case, we aren't sure if it's this PR even or if it's something in Docker on Windows 10. But, if someone else has a chance to test this out, please report back any findings!
Moving along, no other updates there
Any updates you want to mention @terrywbrady on Docker + DSpace: https://wiki.duraspace.org/display/~terrywbrady/DSpace+Docker+and+Cloud+Deployment+Goals

Terry Brady [3:23 PM]
I am taking an AWS course. As my final project, I plan to create a system that will let you select a DSpace PR and run it within docker on AWS. Once I have it completed, I will provide a demo. I'll be happy to turn over the code if anyone is interested in making the service available.
That is my attempt to prototype what is outlined on the wiki page above.
No other updates here.

Tim Donohue [3:25 PM]
Sounds great
Ok, so I think that's it for the usual updates in this DevMtg. Any other topics (from anyone) for today?

James Creel [3:26 PM]
Maybe I'll just mention that there will be a DSpace session at the Texas Conference on Digital Libraries in 11 days

Terry Brady [3:26 PM]
What is the topic?

James Creel [3:26 PM]
Oh, a few things: https://web.cvent.com/event/9c897e8e-c1bf-4ff2-aa26-68682e223059/websitePage:645d57e4-75eb-4769-b2c0-f201a0bfc6ce
web.cvent.com
Agenda - 2019 Texas Conference on Digital Libraries
2019 Texas Conference on Digital Libraries.
It's part of the pre-conference on May 21 from 1:00 to 4:30.

Terry Brady [3:28 PM]
That sounds interesting!
If you have the chance, spread the word about the DSpace Docker stuff.

James Creel [3:28 PM]
Yeah, very locally oriented, but still cool.
I'll be sure to mention that. I wouldn't be surprised if I'm the only DSpacer in Texas to have tried it out.

Tim Donohue [3:30 PM]
Very cool, thanks for the note on this! Definitely report back on anything interesting from the sessions

James Creel [3:30 PM]
Sure thing.

Tim Donohue [3:30 PM]
Any other topics we want to discuss today? Otherwise, we can call this meeting short & give everyone time back in your day
Ok, not hearing anything else. Let's call it a meeting then. If anything else comes up, give a shout on #dev (or wherever is most appropriate)
Thanks all!

James Creel [3:32 PM]
Thank you, Tim.

Terry Brady [3:32 PM]
Have a good week

Mark Wood [3:50 PM]
Sorry I missed the meeting. I have nothing new to report on Solr 7 this time, as I've been buried in a local project that doesn't want to work. (You may have seen my muttering over on #tech-support .)