

DevMtg 2018-09-26

Developers Meeting on Weds, September 26, 2018

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)
 - Please note that all meetings are [publicly logged](#)

Agenda

Quick Reminders

Friendly reminders of upcoming meetings, discussions etc

- [DSpace 7 Working Group \(2016-2023\)](#): Next meeting is tomorrow, Thurs, September 27 at 14:00 UTC
 - **Signup for our next Community Sprint (Oct 1-12): [DSpace 7 Community Sprints](#)**
- [DSpace 7 Entities Working Group \(2018-19\)](#): Next meeting is TBD
 - Yesterday's meeting (including video recording): [2018-09-25 DSpace 7 Entities WG Meeting](#)
- [DSpace Developer Show and Tell Meetings](#): Looking for more meeting ideas!

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 \(2016-2023\)](#))
 - a. [DSpace 7 Development Status spreadsheet](#)
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. [DSpace Release 5.10 Status](#) - Ready to schedule a release date?
4. [DSpace and Docker](#)
 - a. <https://github.com/DSpace/DSpace/pull/2214>
 - b. <https://github.com/DSpace-Labs/DSpace-Docker-Images/pull/59>
 - c. Other Docker PR's
 - i. Codenvy: <https://github.com/DSpace-Labs/DSpace-Docker-Images/pull/54>
 - ii. RDF: <https://github.com/DSpace-Labs/DSpace-Docker-Images/pull/58>
5. Follow-up on "DSpace Top GitHub Contributors" site: <https://tdonohue.github.io/top-contributors/>
 - a. Showed this to others internally (in DuraSpace), and this is likely to turn into a hosted DuraSpace site (covering all projects). Coming soon (hopefully by next week?)
6. Follow-up on Curation Task Reporting ([PR 2180](#))
7. Brainstorms / ideas (*Any quick updates to report?*)
 - a. [Bulk Operations Support Enhancements](#) (from [Mark H. Wood](#))
 - b. [Curation System Needs](#) (from [Terrence W Brady](#))
 - i. [PR 2181](#) implements per-run task parameters. Ready for review.
 - ii. [PR 2180](#) improves reporting. Needs a little more testing
8. 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. 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 `SessionFactory.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

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 [3:00 PM]

@here: It's DSpace DevMtg time. Agenda is at <https://wiki.duraspace.org/display/DSPACE/DevMtg+2018-09-26>
Let's do a quick roll call to see who all is able to join us today

Mark Wood [3:01 PM]

Hi

Terry Brady [3:01 PM]

hello

Tim Donohue [3:02 PM]

Well, it may just be the 3 of us, but we can go ahead and get started
Friendly reminder (mostly to those looking at this later) that our next DSpace 7 Community Sprint is *next week*. Still time to signup at: <https://wiki.duraspace.org/display/DSPACE/DSpace+7+Community+Sprints#DSpace7CommunitySprints-Participants>
We also had our 3rd Entities WG meeting yesterday. Video is at <https://wiki.duraspace.org/display/DSPACE/2018-09-25+DSpace+7+Entities+WG+Meeting> ... and lots of recent discussion on #entities-wg channel too
Beyond that, I don't have any other significant updates to share for DSpace 7 or 6.x. So, I think that about covers topics #1 and 2 on the agenda, unless anyone else has something to add?
not hearing any...topic #3. DSpace 5.10 updates
<https://wiki.duraspace.org/display/DSPACE/DSpace+Release+5.10+Status>
@terrywbrady: my main question here was whether to schedule a date for this release yet? any closer to that?

Terry Brady [3:05 PM]

@pbecker verified that RDF works with Java 7, so I think we have a green light.

Tim Donohue [3:05 PM]

awesome!

Terry Brady [3:05 PM]

How does next Thurs sound?

Will you guys be available if I run into trouble with the release process?

Tim Donohue [3:06 PM]

Next Thurs as in Thus, Oct 4?

Terry Brady [3:06 PM]

correct

Mark Wood [3:06 PM]

I should be available most of the day.

Tim Donohue [3:06 PM]

I'm around. It's during the Community Sprint, but I should still be able to chip in with advice as needed

Terry Brady [3:07 PM]

Perfect. Let's make that the plan.

Tim Donohue [3:07 PM]

Sounds good, Thurs, Oct 4 it is! Thanks, @terrywbrady
Moving along then.. @terrywbrady you are also up for topic #4...and update on DSpace + Docker
<https://wiki.duraspace.org/display/DSPACE/DSpace+and+Docker#DSpaceandDocker-Proposal:SupportMultipleJDKVersions>

Terry Brady [3:09 PM]

As @pbecker and I were testing the RDF stuff, we determined that it would be useful to publish multiple images per DSpace branch in order to certify with different versions of Java.
If you look at the table on the linked page, you will see the versions that I propose we support.
The table also contains 2 tag naming schemes.
Before we talk about tag names, how does the jdk version support look to you all?

Tim Donohue [3:11 PM]

It's fine, though I think it could be slightly simplified by dropping JDK 8 from `dspace-4_x` (and possibly even `dspace-5_x`). Neither of those releases claimed to support JDK8, though I know 5.x actually does (and

many people run it on JDK8)

Mark Wood [3:11 PM]
Looks right.

Terry Brady [3:11 PM]
I will drop it for 4x.

Tim Donohue [3:11 PM]
I'm not against testing these branches against JDK8, but TBH, I have no idea of 4.x even "works" right with JDK8 sounds good

Terry Brady [3:11 PM]
Since it is in common use for 5x, I think it makes sense to keep.
In my minimal docker testing it works, but that has not been very extensive.
Any need for tomcat9 support?

Tim Donohue [3:12 PM]
I'd start small, we can expand later. I don't know of anyone using/testing on Tomcat9

Mark Wood [3:13 PM]
I *have* 9, but do my testing on 8.5

Terry Brady [3:13 PM]
Regarding tagging - proposal 1 always includes the jdk. Proposal 2 only labels the non-primary jdk in the tag name. Your thoughts?

Tim Donohue [3:14 PM]
These tags are going to be in the "DSpace-Docker-Images" repo?

Terry Brady [3:14 PM]
yes

Tim Donohue [3:14 PM]
I'd probably lean slightly towards "proposal 1" as it's more explicit as to which version of JDK is available

Terry Brady [3:15 PM]
@mwood what do you think?

Kim Shepherd [3:15 PM]
hi all

Mark Wood [3:15 PM]
All the same is simpler and you don't have to ask yourself "what is the 'primary' version for this one?"

Terry Brady [3:15 PM]
hi @kshepherd

Mark Wood [3:15 PM]
Hi.

Terry Brady [3:15 PM]
Do you need us to catch you up on the discussion?
It sounds like Proposal 1 is the winner.

Tim Donohue [3:16 PM]
sounds fine

Terry Brady [3:17 PM]
In the table below the one we have looked at, I propose that we also produce docker images optimized for testing: (1)local host restriction removed for solr (2)ssl required removed from rest service.
This would allow someone testing the system to access all web services without overrides.

Kim Shepherd [3:17 PM]
i've scrolled up :slightly_smiling_face:

Terry Brady [3:17 PM]
It would mean creating 2 images for each branch

Tim Donohue [3:19 PM]
Do those two changes warrant creating a new image? Just wondering out loud if those options are so widely

tweaked that we really need a "test" image
(I'm fine either way...just asking to clarify)

Terry Brady [3:20 PM]

When I am testing, I would prefer to have access to those services without having to manually modify files.

Kim Shepherd [3:20 PM]

i can see that that would make it easier for people to test REST.... though for Solr, i personally find it quite easy to just ssh tunnel into the server to get around solr localhost restriction, so that's pretty easy to document as a workaround

Tim Donohue [3:21 PM]

I'm assuming the base images must not have any form of SSL then? So, REST essentially won't work there until they enable SSL?

Terry Brady [3:21 PM]

I think so. I do not know how to add ssl to the images.
That might just be my lack of knowledge. (I had also forgotten about the ssh tunnel option)

Kim Shepherd [3:22 PM]

is the issue around building/configuring a cert in the base image?

Terry Brady [3:23 PM]

Yes

Tim Donohue [3:23 PM]

I guess I'm just trying to understand the key differences / use cases that differentiate these two image types "base" and "test". Why would someone want to use "base"?

Terry Brady [3:23 PM]

Base more closely matches a production build...
(I am excited to get some additional perspectives on this stuff)

Tim Donohue [3:24 PM]

I guess I'm wondering whether we want to even imply we support a "production" build yet. Should we just have a "test" set of images?
I admit, I don't have a strong opinion here... I'm just trying to wrap my mind around why 2 images per branch (edited)

Terry Brady [3:25 PM]

Maybe so... the build would be faster without 2 images
If we make the "test" the default, do we simply override the rest service or do we override rest and solr for convenience?

Tim Donohue [3:26 PM]

If the build is faster, I think that'd make me lean slightly towards calling these .. "for test/demo only" and documenting *why* (cause we disable REST SSL , etc)

Terry Brady [3:27 PM]

I like that idea.

Tim Donohue [3:27 PM]

overriding REST and SOLR is fine by me. I can understand why someone might want both in test/demo environments. We should just document those overrides, obviously

Kim Shepherd [3:28 PM]

i'm happy to contribute a quick "ssh tunnel recipe" for the solr override if you like, for the README - that would also enable users to test the behaviour of the LocalhostRestrictionFilter working as advertised (when visiting directly) but also access Solr via the forwarded local port

Mark Wood [3:28 PM]

I would suppose that someone setting up production would want to customize so much stuff that it would be just as easy to set it up without a container. (Maybe that's just me.) (edited)

Kim Shepherd [3:28 PM]

but i can see how disabling it might also lower barrier of entry for some

Tim Donohue [3:29 PM]

@mwood: yes, they likely could use this as a "starting point", but for production they'd want to tweak it anyhow for local needs, etc.

Terry Brady [3:29 PM]

My hope is for the very low barrier ... I want these images to help encourage new contributors.
OK, in my current plan we do still need to build 2 images for DSpace 5: jdk7 and jdk8.
Eventually, we will support jdk9 for DSpace 7 I assume...

Tim Donohue [3:30 PM]

I'd vote override REST and SOLR, but we document where those overrides exist -- so that if folks wanted to, they can remove them in their own image.

Terry Brady [3:30 PM]

So, now we get to the ugly issue of the automated builds on Docker Hub...

Mark Wood [3:30 PM]

IIRC Java 9 is already dead. 11 just shipped as the next LTS version.

Tim Donohue [3:31 PM]

We should set aside JDK discussions for DSpace 7. That's a whole other discussion, which we don't need to solve in this Docker discussion :wink:

Terry Brady [3:31 PM]

From a command line, you can successfully use a Dockerfile located in a subdirectory. But, if you attempt to do that in the Docker Hub automated build, the directory context switches and the build fails. Therefore, each Dockerfile that we want to use must reside in the root dir.
Thus, you will see my proposed Dockerfile names in the table.

Tim Donohue [3:33 PM]

I presume having multiple Dockerfiles in the same directory doesn't much matter, as you just call which one you want

Terry Brady [3:33 PM]

Exactly.
And I confirmed that I can run multiple auto builds from one branch.
<https://hub.docker.com/r/terrywbrady/dspace/builds/>

Tim Donohue [3:35 PM]

Ok, having multiple is fine by me. We just need to be careful to clean up old ones as we create new branches.
So, when we are done with supporting JDK8, we'd have to cleanup the `Dockerfile.jdk8` from `master`

Terry Brady [3:35 PM]

I will plan to refine <https://github.com/DSpace/DSpace/pull/2214> based on this conversation.

Tim Donohue [3:35 PM]

And it seems like we'd only have 1-2 Dockerfiles per branch anyhow at this time

Terry Brady [3:36 PM]

Once this is approved, I will port to 6x, 5x, and 4x adjusting the tomcat and ant versions as needed.

Tim Donohue [3:36 PM]

This all sounds good to me. Thanks for all your hard work on this @terrywbrady
Is there any other feedback/advice you need here?

Terry Brady [3:37 PM]

I have 2 other bits of Docker news to share

Tim Donohue [3:37 PM]

go for it

Terry Brady [3:38 PM]

I have cleaned up the AIP ingest support files for Docker: <https://github.com/DSpace-Labs/DSpace-Docker-Images/pull/54>
based on your comments @tdonohue.
I will go ahead and merge this later this week unless you all have suggestions.

Tim Donohue [3:38 PM]

awesome! I saw you pinged me, but hadn't had a chance to look at it

Terry Brady [3:39 PM]

@pbecker and I created a compose file that includes an RDF triplestore. <https://github.com/DSpace-Labs/DSpace-Docker-Images/pull/58>
I will also merge that later this week. Comments are welcome.
There may be enough new Docker stuff to show off at a Show and Tell meeting if you all think that would be

interesting to the community.

I am now able to run our published images in Codenvy... so it functions like a Cloud provider running Docker containers.

Tim Donohue [3:40 PM]

@terrywbrady: Tiny thing on PR#54, as that initially included the AIP Zips in the PR, I'd recommend *squashing* those commits...that way the AIPs themselves never even make it into the Git history

Terry Brady [3:41 PM]

Will do.

Tim Donohue [3:41 PM]

You could also just do a "squash" merge (which is a merge option)

Terry Brady [3:41 PM]

I have lots of squashing to do after my testing the last 2 days...

Tim Donohue [3:42 PM]

I think a Docker update at a new Show & Tell would be great. It seems like we are getting down to the area of "finalizing the setup" and promoting/creating basic "best practices" that would be worth sharing/promoting Plus the past Docker Show & Tells were well attended...so this seems to be an area of interest to keep sharing /updating folks on

Terry Brady [3:42 PM]

I hope I can get the 3 of you converted to using it in your testing! I will look for a time on the calendar.

Mark Wood [3:43 PM]

Oh for an option "squash commit1..commit2".

Tim Donohue [3:43 PM]

I have a plan to plan to use Docker (in a VM environment). I just need to find "free time" to move in that direction.

Terry Brady [3:43 PM]

Thanks for letting me monopolize the meeting!

Tim Donohue [3:44 PM]

Ok, is that it for our Docker topic then? :wink: It's definitely great seeing this move along so much week by week

Terry Brady [3:45 PM]

I got what I needed. Thanks

Tim Donohue [3:45 PM]

thank you as well

So, moving along to topic #5... a followup on my end regarding "DSpace Top GitHub Contributors"

As noted last week, I have this all "working" at <https://tdonohue.github.io/top-contributors/> (code linked from this homepage)

tdonohue.github.io

DSpace's Top GitHub Contributors

Ranking DSpace GitHub contributors, since 2018

And in the past week, as I've showed it off internally (in DuraSpace), it's got other folks asking for it to be DuraSpace-wide :slightly_smiling_face:

Terry Brady [3:47 PM]

My boss shared the link with folks in our library, so I appreciated it!

Tim Donohue [3:47 PM]

So, before an official "launch" / announcement, I'm working on making this thing cross-project. It soon will have DSpace, Fedora & VIVO stats..

Terry Brady [3:47 PM]

@tdonohue you could answer my StackOverflow question related to this. <https://stackoverflow.com/questions/51972020/github-graphql-query-count-pr-reviews-by-user-for-a-given-month>

Tim Donohue [3:47 PM]

And that also likely means it will eventually live at "duraspace.github.io" in the near future (hopefully as early as next week sometime)

@terrywbrady: yes, I can answer that SO question once this thing has a final home...as I suspect links to it would help

So, that's the update I wanted to share. It's still in "coming soon" status, but the stats that are there are

all accurate for DSpace. I'll keep you posted on once this has an official launch pending
Any questions/comments before moving along to next topic?
Not hearing any...and 10 mins left, so we should have time for one more topic here :slightly_smiling_face:
Topic #6, following up on Curation Task reporting in <https://github.com/DSpace/DSpace/pull/2180>
@mwood: did you want to update us here?

Mark Wood [3:51 PM]
I still haven't tested reporting from workflow task attachments.

Tim Donohue [3:52 PM]
Is there any feedback/comments you still need from us? I see there are a few new commits since we talked last week

Kim Shepherd [3:53 PM]
i can help test that, looks good

Mark Wood [3:53 PM]
Well, I split off the backend report writing into plugins: one does files, the other sends output to dspace.log.

Kim Shepherd [3:53 PM]
and ooh, unit tests :slightly_smiling_face:

Mark Wood [3:53 PM]
Thanks for testing.
The file writer no longer writes to [DSpace]/log

Tim Donohue [3:54 PM]
yes, testing would be good...thanks @kshepherd for offering
Ok, well, if there's no feedback/discussion needed, we can always just check back in on this next week (or whenever you are ready). In the meantime, testing / more eyes are obviously welcome

Mark Wood [3:56 PM]
Well, the usual feedback is needed: "is this what we want?"
I think there is some unresolved debate over how fancy the basic report output should be?

Tim Donohue [3:57 PM]
I'm hoping @kshepherd (with fresh eyes) might give some perspective on that question. I know we talked around that a bit last week. It *sounds* like a reasonable idea to me, but I admit, I haven't tested it or figured out how I'd use it personally

Kim Shepherd [3:58 PM]
i use and have used curation tasks for a bunch of things and i've often wanted improved reporting/logging (particularly simple structured reports), so i'll put myself back in those shoes

Tim Donohue [3:58 PM]
thanks!

Mark Wood [3:59 PM]
It is now fairly simple to write your own serializer, if you need something special, but it has to live with the Appendable interface.
BTW passing an object rather than a path string to Curator made testing a snap. (edited)

Tim Donohue [4:00 PM]
So, as we are nearing the top of the hour, it sounds like we should touch back on this topic over the next few weeks (and in the meantime add comments/thoughts into the PR too).

Mark Wood [4:01 PM]
Yes please.
There's also the related "run parameters" PR if anyone is interested.

Terry Brady [4:01 PM]
Have a good week. I will drop a note in our docker channel that folks might want to review the meeting logs.

Tim Donohue [4:01 PM]
Unfortunately, I have to run here quite shortly. So, let's wrap up today's meeting (no new topics).
Additional discussion is always welcome in this channel or #dev

Terry Brady [4:02 PM]
@mwood, I apologize that I have not yet had a chance to test these PR's

Mark Wood [4:02 PM]

Yes, I have to go now too.

Kim Shepherd [4:02 PM]

cheers all

Tim Donohue [4:02 PM]

Thanks again all for the good discussion & PRs. Lots of interesting work going on!

Mark Wood [4:02 PM]

Thanks, all!

Kim Shepherd [4:02 PM]

coffee time