2023-06-01 Product Council Meeting notes

Date
01 Jun 2023

Attendees

Recording
https://recordings.openlibraryfoundation.org/folio/product-council/2023-06-01T09:30/

Discussion items

<table>
<thead>
<tr>
<th>Time</th>
<th>Item</th>
<th>Who</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 min</td>
<td>Announcements</td>
<td>All</td>
<td>WOLFcon early bird registration open until June 18th. Call for proposals is open until June 5th; aiming to have an agenda ready before the early bird registration closes. Draft proposals are okay; they can be revised later. FOLIO Council election results should be coming out soon after the Community Council meeting on Monday.</td>
</tr>
</tbody>
</table>
This comes out of various documents and the meeting in Stanford, and includes two reports from EBSCO.

Problem statement: we don't have a reliable and performant solution for loading records to support daily workflows and monthly/quarterly workflows. (Leaving aside system migrations for the moment.) Record loads of less than 1000 records can timeout. Libraries are unable to manage basic workflows such as loading electronic records. FOLIO users must plan to work off-hours to complete import jobs.

Slide 6 describes in-progress efforts across these topics: architecture, infrastructure, development, and product management.

"Large" loads are about 100,000 records, and there need to be plans to address loads of that size and much smaller sizes. Also noting that some types of loads are more performant than other types, so it isn't necessarily about the number of records. "Chunking" can also allow for smaller loads to run interspersed with a larger load. (Single record imports jump the queue right now.)

In addition to the documentation prepared, interviews with current users of FOLIO help make the issues more understandable. (There is a governance question of who "customer" or user is, and if it is inclusive of all hosting providers and those not using a hosting provider.)

Performance is one issue, as is the capabilities of the logic that is a part of data import.

Questions from the meeting include how bugs are prioritized and whether there can be a dashboard of work that is happening on Data Import. (A dashboard is in the works.) Libraries need to be able to know when the functionality will be implemented so they can plan for their local workloads, including when functionality is deemed out-of-scope.

Short term: complete work in progress and address critical production needs (continue performance improvements and reliability/scalability improvements). Collaborate with other hosting providers and self-hosted institutions. Provide realistic performance benchmarks. Define an approach for addressing failed records and logging issues.

Mid-term: review the data import roadmap; continue architecture and infrastructure improvements; development to address failed records and logging issues.

Longer-term (12 months): continue architecture, infrastructure, and development improvements

Proposal for release changes: extend the Poppy release to the current Q-release date (November 2023). There is functionality (other than data import) that is not ready that is driving this decision.

PPT - Data_Import_Improvements.pptx

Further discussion to happen over Slack, and aiming for a decision to be made on Monday.
00:03:35 Tod Olson: I think also “good middle of the night” for a few people.
00:05:08 Kristin Martin: Earlybird through June 18
00:08:15 Kristin Martin: https://wiki.folio.org/display/PC/2023-06-01+Product+Council+Meeting+notes
00:14:04 Christopher Spalding (EBSCO): Kristin, weâ€™ll send a link to this deck so it can be added to the documentation.
00:14:36 Kristin Martin: Reacted to “Kristin, weâ€™ll send...” with δŶ.
00:31:21 Charlotte Whitt: Those P1/P2 bug tickets are they all tagged with the label ‘support’ - and monitored by the Support SIG?
00:31:48 Erin Nettifee: Related - is there a jira dashboard for this work in progress that could be shared?
00:32:09 Charlotte Whitt: Reacted to “Related - is there a...” with δŶ.
00:33:56 Erin Nettifee: Yes - a board would help with transparency of the work
00:36:03 Gang Zhou | SHL: Reacted to “Yes - a board would...” with δŶ.
00:36:05 Jennifer Eustis: Reacted to “Related - is there a...” with δŶ.
00:36:25 Jenn Colt: It does. I thought it was already there.
00:36:49 Jennifer Eustis: I thought it was already there as well. It's to allow the single record imports/overlays to go through
00:37:29 Jenn Colt: Closed https://issues.folio.org/browse/MODSOURMAN-662
00:37:54 Felix Hemme: Why can't multiple jobs run in parallel?
00:38:24 Mark Veksler: Replying to “I thought it was air...” partially. I believe it depends from where import is launched.
00:38:38 Felix Hemme: Replying to “Why can't multiple j...”

Or at least single Import Independent from importing MARC files?
00:38:44 Jenn Colt: We have provided such stats repeatedly. The Di wiki page gas suggested standards with large goal = 100k
00:39:06 Lloyd (He/Him): Replying to “Why can't multiple j...”

Or fast enough that it appears to be parallel.
00:39:14 Jenn Colt: Replying to “We have provided suc...” Has...
00:39:14 Jennifer Eustis: Reacted to “We have provided suc...” with δŶ.
00:39:31 Christie Thomas (she/her): Replying to “Why can't multiple j...”

This is happening for us in Nolana. Single record imports will process while a data import file is also processing.
00:39:49 Felix Hemme: Reacted to “This is happening fo...” with δŶ δŶ
00:40:29 Jennifer Eustis: That would be great.
00:40:45 Laura Daniels: Reacted to “That would be great.” with â±
00:40:50 Colin V. (he/him): Reacted to “That would be great.” with â±.
00:40:50 Lynn Gullickson Spencer: Reacted to “That would be great.” with â±.
00:40:53 Erin Nettifee: Is there a document anywhere that lists the requirements for data import and what has been implemented and what hasn't? I'm aware of https://wiki.folio.org/display/FOLIOtips/Data+Import+Functionality+Implementation+Table but it is so high-level, and where the issues really come out is in the nitty gritty of what data import needs to be able to do.
00:41:08 Lloyd (He/Him): 100k is the most I would want to do at one time anyway. More would choke my discovery system.
00:41:11 Christopher Spalding (EBSCO): Replying to “Why can't multiple j...”

@Lloyd (He/Him) The design approach for chunking would allow for parallel processing.
00:41:26 Jenn Colt: I donâ€™t want to drag us into a rabbit hole. I just wanted to say itâ€™s been discussed.
00:41:37 Erin Nettifee: e.g., “Update SRS MARC BIB via Data Import” is so general, and doesn’t help people understand what mapping is possible, what functions are supported under that umbrella, etc.
00:41:40 Christie Thomas (she/her): And large during the work day / peak hours is different from large during off hours.
00:41:56 Erin Nettifee: I think we have a very weak understanding across the board of what the app is actually supposed to be doing.
00:42:42 Erin Nettifee: I helped with documentation discussions around DI, and there are so many instances of “Wait, I thought that was implemented?” â€¢ concrete requirements that could be referenced would be so helpful. Even if we have to go back and drag through years of development work to articulate it.  
00:43:00 Tod Olson: So “large” is a function of both record set size and complexity of import profile?
00:43:25 Jenn Colt: Replying to “I helped with docume...”

Part of that is because these discussions start over and over with information degradation.
00:43:35 Jenn Colt: Replying to “I helped with docume...”

Fractal telephone
00:43:37 Christopher Spalding (EBSCO): Reacted to “So “large” is a func...” with δŶ δŶ
00:43:47 Tod Olson: Reacted to “Fractal telephone” with δŶ.
00:44:11 Christie Thomas (she/her): Replying to “So “large” is a func...”

Yes, and also time - 100,00k is small for an off hours import. We should be able to do larger files during off hours.
00:44:44 Christie Thomas (she/her): Replying to “So “large” is a func...”

100,000
00:45:11 Erin Nettifee: Replying to “I helped with docume...”

Not disagreeing. But the requirements that are supposed to be in place need to be written down. I'm asking, are they?
00:45:13 Charlotte Whitt: Reacted to “And large during the...” with δŶ δŶ.
00:46:58 Jenn Colt: Replying to “I helped with docume...”

Dev has not been linear. Something is resolved and then breaks, repeatedly. What is its current state? Who knows?
00:47:06 Lloyd (He/Him): Replying to “So “large” is a func...”

That gets into scheduling. It would be fine if I could schedule several 100k loads for off hours.
00:47:39 Jenn Colt: Replying to “I helped with docume...”
I really need to push back on this being a SME problem

00:50:40 Charlotte Whitt: The paper Corrie Hutchinson has written up, reflects customers expectations.

00:50:52 Lynn Gullickson Spencer: Reacted to "The paper Corrie Hut..." with ðŸ’ðŸ»

00:51:15 Charlotte Whitt: Will these customer interviews/feedback build upon this document, or start the process all over again
00:52:22 Thomas Trutt: I would suggest starting over, clean slate of expectations, and see how the current system address those needs and where the gaps are. Otherwise you get into the issue of this needs fixed instead of requirements.

00:52:47 Jenn Colt: Can we please stop and define customers?

00:55:25 Erin Nettifee: Replying to "I would suggest star..."

That is what is missing - the list of requirements that the app is supposed to meet. Without that, it is very confusing when someone is looking at a thing that doesn't work. You can't know where to start if you don't know if the app is supposed to be able to do the thing that didn't work.

00:56:26 Thomas Trutt: Reacted to "That is what is miss..." with ðŸ’ðŸ»
00:56:33 Corrie Hutchinson: Reacted to "That is what is miss..." with ðŸ’ðŸ»
00:57:48 Thomas Trutt: But that flow logic sounds like a base requirement.

This is the theory behind the homegrown tools being developed, imo
01:00:21 Jenn Colt: Burn it down is always appealing in theory

01:01:50 Aleksey Petrenko: Reacted to "https://wiki.folio.o..." with ðŸ’
01:01:54 Thomas Trutt: ðŸ™’ Would it make sense to simplify and streamline DI and allow for plugins, that add functionality, instead of trying to put all the features into one container?? (Just throwing thoughts out there..)
01:01:59 Thomas Trutt: Reacted to "Burn it down is alwa..." with ðŸ”ƒ
01:03:09 Jenn Colt: Replying to "ðŸ™’ Would it make sen..."

Sometimes starting over saves time and resources compared to patching bad code, or a bad idea.
01:03:47 Thomas Trutt: Replying to "Burn it down is alwa..."

Owen Stephens: I think itâ€™s fair to say that bug prioritisation is an art rather than a science. The PO (usually the PO at least) has to balance many factors when prioritising bugs and scheduling work

01:04:37 Jenn Colt: Sharing the link would be wonderful
01:05:05 Stewart Engart, Ph.D.: Reacted to "That is what is miss..." with ðŸ’ðŸ»
01:05:26 Anya : Also please come to the support sig - every Monday
01:05:41 Stewart Engart, Ph.D.: Reacted to "Also please come to ..." with ðŸ‡‰
01:06:14 Steph Buck: Reacted to "I think itâ€™s fair to..." with ðŸ’ðŸ»
01:10:17 Jenn Colt: Sharing the link would be wonderful
01:10:56 Alexis Manheim: Reacted to "Sharing the link wou..." with âž•
01:11:03 Erin Nettifee: Replying to "Burn it down is alwa..."
i agree with Khalilah that we don't know enough yet to know if that has to happen. we need to stay up in the workflow / functionality stage and let the developers figure out if something can or can't be done.

01:12:18 Lisa McColl: Yes - nice to see the "Limited functionality" row.
01:12:35 Charlotte Whitt: Reacted to "Yes - nice to see th..." with δÝδÝ
01:12:40 Christie Thomas (she/her): Reacted to "Yes - nice to see th..." with δÝδÝ
01:12:52 Charlotte Whitt: Reacted to "Sharing the link wou..." with åž•
01:15:17 Thomas Trutt: If other institutions are using external tools is it worth while to look at what they are doing and adopt that?
01:16:42 Charlotte Whitt: The German community uses mod-inventory-update
01:16:48 Owen Stephens: Who is åœ€œ€œ€ in that context? åœ€™m rather unclear what the PC can do beyond banging the drum about the issue?
01:16:54 Erin Nettifee: But that doesn't support MARC, right Charlotte?
01:17:11 Khalilah Gambrell (EBSCO): We = Community
01:17:42 Charlotte Whitt: https://github.com/folio-org/mod-inventory-update
01:18:31 Charlotte Whitt: Can support MARC and any format - but will require setting up a the data flow
01:19:42 Charlotte Whitt: Reacted to "But that doesn't sup..." with å®
01:20:29 Aaron Neslin: +1 - I appreciate the steps that have been taken, but we've been pointing out these issues for literally years
01:21:01 Jenn Colt: Reacted to "+1 - I appreciate th..." with δÝ
01:21:02 Thomas Trutt: Very nicely put Jenn...
01:21:08 Erin Nettifee: Reacted to "+1 - I appreciate th..." with δÝ
01:21:39 Jenn Colt: Understanding how you will know would be a great step. What is the discovery? And how is it different from the. Last two years of discovery?
01:24:46 Anya : University of the Arts is now the newest library to go live on Full FOLIO
01:25:01 Jenn Colt: I watched Picard over the last week and was reminded of DI: åœ€œ€œ€ always comes later than we think it should.å€ https://en.wikiquote.org/wiki/Jean-Luc_Picard
01:25:12 Anya : Reacted to "I watched Picard ove€" with δÝ
01:25:26 Owen Stephens: Replying to "University of the Aræ€" Is that the UK based one?
01:26:14 Charlotte Whitt: Reacted to "I watched Picard ove..." with δÝÇ€
01:26:23 Erin Nettifee: Reacted to "Can support MARC and..." with δÝ
01:26:49 Anya : Reacting to "University of the Aræ€"
No - PA,USA
01:27:39 Thomas Trutt: Reacted to "I watched Picard ove..." with åœ€, åœ€
01:28:01 Owen Stephens: Reacted to "No - PA,USA" with δÝÇ€
01:28:42 Mark Veksler: Reacted to "University of the Ar..." with δÝ
01:28:45 Owen Stephens: This is a big decision to make so late in the release cycle
01:29:45 Charlotte Whitt: Reacted to "This is a big decisi..." with δÝδÝ
01:30:02 Julie Bickle: +1 Owen
01:30:53 Martina Schildt | VZG: Reacted to "+1 Owen" with δÝ
01:31:37 Martina Schildt | VZG: Completely agree Owen
01:31:48 Thomas Trutt: In general I think less releases per year is a good idea. The time spent on Bugfest and squashing those bugs could be put into more dev cycles.
01:32:22 Jennifer Eustis: Reacted to "In general I think l..." with δÝ
01:33:36 Charlotte Whitt: Some libraries have postponed upgrade to Orchid, and then planned to do Poppy
01:34:03 Charlotte Whitt: Reacted to "In general I think l..." with δÝ
01:35:20 Owen Stephens: I will also note that our development plans would have been very different if we were not planning to meet the poppy deadline
01:37:50 Charlotte Whitt: Similar for the work we do for Mainz (Gutenberg) and Hebis (Odin)
01:37:58 Charlotte Whitt: Reacted to "I will also note tha..." with δÝδÝ
01:38:02 Owen Stephens: Reacted to "Similar for the workæ€" with δÝ
01:38:42 Martina Schildt | VZG: Reacted to "I will also note tha..." with δÝδÝ
01:38:43 Martina Schildt | VZG: Reacted to "Similar for the work..." with δÝ
01:39:50 Owen Stephens: Who will make the final call and when?
01:40:00 Wayne Schneider: @harry that is a little more complicated than you are suggesting due to dependencies in both FOLIO interfaces and UI libraries.
01:40:11 Harry: Reacted to @harry that is a lit..." with δÝ
01:40:37 Owen Stephens: And the new CSP does not allow for non P1/P2 to be addressed via releases
01:40:40 Kirstin Kemmer-Heek: Reacted to "And the new CSP does..." with δÝ
01:40:48 Charlotte Whitt: Maybe we could do a Poppy Service Patch - not just bugs, but real work too
01:40:59 Martina Schildt | VZG: Reacted to "Maybe we could do a ..." with δÝ
01:42:57 Jennifer Eustis: thank you everyone