

Theme: Shanghai library FOLIO project

Time: May 25, 2021 07:00pm (EST) / May 26, 2021 07:00am (GMT+8)

Attendees:

Vincent Bateau (Enterprise Architect, EBSCO)
Gang Zhou (Project manager, Shanghai library)
Sha Jiang (Technical Director, Jiatao)
Lucy Liu (Product Owner, Folio China)

Notes:

1. ElasticSearch

Resources shared before and in the meeting:

[Search](#)

[Inventory Search - POC Overview](#)

TC meeting notes: [2021-04-28](#); [2021-05-05](#); [2021-05-12](#)

Vince:

- We have completed the work planned for R1 for mod-search. There were statements in the earlier meetings that mod-search was not part of FOLIO. But it is.
- Mostly the deliverable is the back-end work. It consists of two pieces:
 - One is the standing up ElasticSearch and having it indexes and so forth. Against that we put mod-search, which is a module that queries the index that has been created. It also supports the CQL syntax that the folio project has adopted. So there is a parser included in it. And it supports a number of things/features (keyword search, facets, etc.) .
 - The other piece is that we need to feed data into the index so that we can find search results. To that effect, there were changes to mod-inventory-storage. They were made and completed. The purpose is to capture certain events as they occur and push event messages onto the kafka queue that we use, on the right topic. The kafka queue is read by a corresponding subscriber on the ES side and updates the index with the information provided. That was included in the mod-inventory-story module.

Both modules are part of the Iris release.

- Those are backend pieces that are meant to be used to provide functionality. But it's difficult for people to evaluate what is this doing, how is it better than the existing search. So the team produced a bonus deliverable, which was an alternative UI (ui-inventory-es) whose sole purpose was to exercise the mod-search apis. It sits alongside. You can have both of them run in folio.
- Now the community is at the point to decide what they want to do next.
 - They need to decide whether they want to modify UI search to make available the changes that were put into the ui-inventory-es so that the ui-inventory would support the new elasticsearch capabilities. That hasn't been decided yet.

- The community has to decide whether they want to completely move to ES or offer both options just in case the hosting institution does not want to run ES. So they could have a switch to use one or use the other. That's where they are at in terms of where they are going to go.
- There was some testing done on the ES. It clearly demonstrated it was much more reliable and effective. It produced faster and better results. It also produced exact counts on the results which the current Postgre does not support.

Sha Jiang: I understand that mod-inventory-storage is now connected to the ES. Do we have further plans for other storage modules?

Vince: Not yet. Right now the search capability is provided as part of the inventory. The inventory sends the messages into Kafka. This is another decision the community has to make: if they want to do an inventory search or if they want to do a complete ES search. What else is going to benefit from using elasticsearch? The mechanism is there. That's an option as well. Inventory was the biggest problem to solve, though.

Sha Jiang: Did we test the efficiency of the changes?

Vince: Yes. Please refer to this wiki page <https://wiki.folio.org/display/DD/Inventory+Search+-+POC+Overview>.

Sha Jiang: What's the future plan for ElasticSearch in the community?

Vince:

- The backend modules are in the release. The community has to decide what to do with that. The first decision they need to make is what to do with the search inventory.
- Some discussions in the meetings: Is this the right search experience? How to make the search dialog more interesting? We prefer to use Solr. etc. The Metadata Management SIG were happy with the ES research.
- My guess is the community will now merge the search UI that is existing in the two different modules into mod-inventory. The only question to be decided is whether we continue to support search on Postgres or not. Whether we have two searches or just one for ES. That will probably be put on the roadmap once the decision is made.
- Iris release has all the necessary pieces to make it work. It's just a matter of making the UI that calls the search service. There is an example of how to do that with the module that was created mod-inventory-es.

Gang Zhou: Any plan to support other apps, such as Users, etc.?

Vince: No plan right now to support Users and other apps. There are ideas of doing it. But the plans have not been formulated. Once we complete the interface, the plan will probably be made up.

Gang Zhou: Any features to support sorting the results by call number?

Vince: Not sure. There were some discussions of extending some of the sort changes. But I didn't see an obvious plan.

Gang Zhou: Is Magda Zacharska the PO in charge of ES? Should we direct questions to her?

Vince: Magda was the PO for the effort that was completed with the POC. There is no PO now because there is no current activity around ElasticSearch. Magda might be the PO for future work. She is knowledgeable and qualified to answer questions about the ES.

2. Data Synchronization

Resources shared before and in the meeting:

[Data synchronization | dependencies across apps](#)
[Automated patron blocks data synchronization](#)

Gang Zhou: In what scenarios will this feature be used? To share data between different apps or to exchange data?

Vince:

- The question is about the general direction. There are two answers to that.
 - On the one hand, the App Interaction SIG produced this document <https://wiki.folio.org/display/AppInt/Data+synchronization+%7C++dependencies+across+apps>. They have some ideas about what we should do with data synchronization.
 - On the other hand, on the technical side, the direction is to basically shift more into an event driven approach.
 - We don't want to have a bulk update. We don't want to have apps that directly call other apis because that causes dependencies and entangling. We want to have something similar to what we did to the ES integration. So we get events sent over from inventory when there are changes to the inventory. They go to the kafka queue. And they are subscribed to and consumed by other apps that need to synchronize. That's the model going forward.
 - This model has been applied a little bit in other parts of the system. [Automated patron blocks data synchronization](#) applies this model. The same model/pattern was applied. The difference is that one uses kafka directly, the other - mod-pubsub. So the direction going forward is to make kafka directly available to ask. The old way of doing it with mod-pubsub is okay, but it has limitations with regard to volume and capacity. It's not going to be able to work for very high volume or high rates of data.
 - The recent efforts we are making is to finalize the availability of kafka as a general purpose utility. That includes setting up the security model that we will provide for kafka, better documentation on setup, and better integration into the deployments and such. The direction is to use an event driven mechanism.

Gang Zhou: Is this in discussion for data synchronization? Or, will there be a feature in the future?

Vince:

- There's active work going on with this, not just discussion. The work is two parts.
 - The first part is that we add new capabilities/new functional features to the platform. They will be added using this model of data synchronization. They will work by pushing events to the message queue.
 - The second effort is to refactor some of the existing applications to change how they work. That's a bigger effort because it's always harder to change something that's already been done. But in the long-term, I think that's going to happen.
- Some of the new stuff we are doing in this way
 - the remote storage work
 - data import: data synchronization between SRS and inventory
 - maybe acquisition in the future

3. Concurrency in FOLIO circulation App

Gang Zhou: We found that when more staff clients used the circulation modules, the response time slowed down. Do you have any reference size?

Vince:

- The question is how many simultaneous users can be supported. Your measurements might be the first that I've heard about. We can ask Chalmers and other live users how they are operating in terms of simultaneous users. Will communicate through the Slack channels.
- Some resources from the Performance Task Force:
 - <https://wiki.folio.org/pages/viewpage.action?pageId=36580308>
 - <https://wiki.folio.org/pages/viewpage.action?pageId=50298970>

Just randomly picked these two. There are more sister pages.

4. Updates of Folio Project at SHL

Gang Zhou:

- We will go live with the circulation module in the open stack next week.
- If we have performance issues. Whom should we contact?

Lucy: Find the PO on [this wiki page](#).

5. PC & TC elections

- Gang Zhou will run for PC.
- Sha Jiang will run for TC.
- It's not decided yet if Vince will run for TC. But Vince will attend the TC meetings anyway.