Theme: Shanghai library FOLIO project

Time: July 14, 2020 07:00pm (EST); July 15, 2020 07:00am (UTC+8)

Link to Zoom meeting

https://zoom.us/j/98935865023?pwd=S0tQL25VWGhUMk5NbWRnaXJDTkdGUT09

Attendees:

Vincent Bareau (Enterprise Architect, EBSCO)

Gang Zhou (Project manager, Shanghai library)

Sha Jiang (Technical Director, Jiatu)

Lucy Liu (Product Owner, Folio China)

Notes:

1. **Feedback to the Search Engine and Circulation group discussions**

   Jiang Sha:
   - Circulation redesign. The company agreed to send developer(s) to join the Core-functional team to help the community improve code.
   - Search engine: We had detailed discussions about the design in the group. The community design was good. What to do next? Should we implement it or wait for others to do it?

   Vince:
   - We can discuss the next steps in Wednesday's meeting. Vince will ask **Mikhail Fokanov** to take feedback and modify his design to establish a blueprint for implementation. The community currently has no development plan for the search engine. The Shanghai team can go ahead to develop it if Shanghai Library needs it urgently. It won't be a problem if we all work on the same blueprint. We may need to know the roadmap/milestones of Shanghai's development work so that the community can review the progress and maybe send people from inventory, index, etc. to provide support. So the direction will be to establish a blueprint, develop (community or Shanghai, no matter who goes first), then collaborate.

2. **Circulation & performance**

   1) **Shanghai team's plan**

      Zhou Gang:

- Confirmed that the team would work with the community to improve performance and would use the patch as a backup.
- The team will redo the performance test with 4-5 servers. The team is preparing the proposal and will share with Vince for review.

2) Expectations on future performance enhancements

Vince: Jakub mentioned there might be more improvement on Okapi. Is Shanghai's choice of Okapi clustering dependent on this, or it's ok with the existing/current performance with Okapi?

Jiang Sha: More performance is good.

Vince: Not sure when the Core-platform will deliver more improvements. Maybe next release (Q3 2020). Just want to make sure it's not a blocker for Shanghai.

3) Collaboration on solving performance issues

Jiang Sha: If we send developers to Core-functional, what are we going to do with Okapi? What projects shall we work on? Do we just pick issues from JIRA, develop and submit code?

Zhou Gang: Do you have an issue list from the Core-functional team? Our concern is the performance issue. Can we get a list of performance issues so that we can estimate its importance to our project?

Vince:
- Given the discussions we've had about rewriting the circulation app and the analysis Taras has done that we have reviewed, the plan is that we can apply some of the changes/recommendations Taras has made and add those changes to the existing circulation app. These will be new issues/tasks in JIRA. The purpose will be to make the circulation app more performant. It's not a complete rewriting, but applying Taras' recommendations to restructure the code.
- The developers will participate in discussions, look at issues available in Jira, assign issues, pull issues and choose issues they will develop for the next sprint and discuss with other team members to find best solutions. The Core-functional team will work together to decide which issues to work on. The Core-functional team has their own meetings and maybe also has their own space on wiki.
- Taras will work as the solution architect with that team. Marc Johnson is the tech lead of the Core-functional team.
- Do you have someone in your mind that will join Core-functional?

Jiang Sha: Yes, already picked some from Jiang Sha's team. Need to check time for the meeting.

Vince: Most of the Core-functional team members are in Europe. (Taras - Ukraine; Marc - UK) Not too bad.

Lucy: How many developers are needed for this project? Will they work full-time or part-time? How long will this collaboration last?

Jiang Sha: The more developers, the better. A group of developers will work part-time.

Vince: Vince's guess is at least one quarter.

Lucy: In addition to sending developers to Core-functional, another way to collaborate is to assign Jira issues directly to the Shanghai team? Are we going to do both? Or one way is sufficient?

Vince: Not both. It's better to work with the team together to avoid confusion.

3. Codex & SRS

Zhou Gang:

- We want to know more about the SRS data model. We are planning to migrate from the current library system to folio. We found the performance of SRS data conversion not good.
- There have been very few updates on Codex recently. Any plan or roadmap for the Codex function in the future?

Vince:

- SRS:
    - SRS was designed for import purposes, not for migration. There have been known performance issues when dealing with large amounts of data. A group of people in the community have been working on that in the past few months. An effort made recently is to work with Texas A&M developers to implement recommendations for how to improve SRS.
    - When SRS does import, the source SRM (Source Record Manager) is responsible for coordinating the import and imports one record at a time. It supports MARC records well now. It takes in a batch of MARC records and applies the workflow. The reason for the slow performance is that the initial implementation focused on complex workflows where it takes the MARC records and then makes multiple calls to the inventory to figure out what has to happen and take different paths.

- ○ Texas A&M is doing high-speed import on a simple case and is trying to figure out an optimization of batching. The approach they take is to use pub-sub. They are adding optimizations to pub-sub to make pub-sub more robust and adding confirmation of delivery to pub-sub. This is the direction.
- Codex
  - ○ Codex was discussed in the design of folio because it can search across physical and e-resources. Codex plays the role of federated search. This concept is extensible to multiple inventory or multiple resources, such as archival collection, institutional repositories, order systems, etc. Codex can also serve as a connection between related resources.
  - ○ The first part of Codex is federated search. It requires all contributors to provide a standardized representation, like Dublin Core. It was never implemented because the community didn't reach an agreement on this.
  - ○ There has been no active development on Codex for a long time. When we introduce ElasticSearch, we can bring Codex by combining indexes from the different modules in folio. So we can use the search engine to implement a better Codex search in the future. ElasticSearch is another to improve Codex.

Zhou Gang: Shall we include Codex in the ES discussion in the future? Shall we combine these two concepts or models?

Vince: Maybe. We need to agree on the simple problems we want to solve first, i.e. to build ES on inventory right now. After that, we can discuss how to create Codex or combine indexes from multiple inventory.