

Theme: Shanghai library FOLIO project

Time: October 28, 2020 07:00pm (EST) / October 27, 2020 07:00am (GMT+8)

Attendees:

Vincent Bateau (Enterprise Architect, EBSCO)

Gang Zhou (Project manager, Shanghai library)

Sha Jiang (Technical Director, Jiayu)

Lucy Liu (Product Owner, Folio China)

Notes:

1. JDK8 upgrade - Some modules require jdk11, do we have a plan for JDK versions in the community?

- Lucy shared a wiki page <https://wiki.folio.org/display/DD/Upgrade+to+JDK+11>. It's said on this page that the Spitfire team had volunteered to test the upgrade to JDK11 and would come up with a step-by-step guide for the other teams to follow. Lucy reached out to **@Khalilah Gambrell**, PO of Spitfire, to get updates. Khalilah shared the Jira Epic# (<https://issues.folio.org/browse/FOLREL-405>), including modules that have upgraded, and the guide (<https://github.com/folio-org/raml-module-builder/blob/master/doc/upgrading.md#version-310>).
- Vince's comments to the above linked resources: In summary, these indicate the following:
 - ❑ Okapi and RMB have both been updated to support JDK11.
 - ❑ For any module built from RMB it is then sufficient to upgrade it to use RMB v. 30.
 - ❑ All RMB modules have now been upgraded to JDK11 - with one exception: mod-login-saml - because it had outdated dependencies that are not compatible with JDK11.
- Sha Jiang: Why should we upgrade to JDK11?
Vince: Because the community has chosen to stick with LTS long-term support versions of jdk. JDK8 is now falling out of long-term support.

2. Does the community have any recommendations for the tools or solutions to use in PostgreSQL and Docker management?

- Vince: Different institutions have different solutions to Docker management. The community is focusing on using kubernetes for container management and, in addition to kubernetes, is looking in many cases to use Rancher. So kubernetes and

Rancher are the directions the community is heading towards. This is similar to what Texas A&M does. The developer's environment is also using this stack. EBSCO is using AWS and ECS. In the future, we will also move to AWS and EKS.

- Gang Zhou: Are there documents or guides we can follow?

Vince: There are some documentations. If you look it up in the SysOp SIG, there's an operator's guide describing TAMU practice. Jason Root from TAMU and Wayne Schnider from Index Data have worked on a document, which is a system administrator's guide describing how to set it up. The document is not finalized. Another place is going to be for the developer environment system in place with the last release. I can find some links and post back to Slack.

- Sha Jiang: A question about memory management for RMB. In docker image or docker instance, if we have many tenants, maybe 1000, for every tenant, RMB would create a connection object, or Postgres circle client object, and cache into memory. If we put all this in docker, we could have many instances. If we have 1000 tenants, the connection objects will consume a large amount of memory. Is that true? How to solve the problem?

Vince: You have one docker container for every module and that module will create a connection for RMB. One thousand tenants in one docker is too much. You could spin out multiple containers with fewer tenants each and load balance between them.

- Sha Jiang: Is there a strategy that we can have a limited number of tenants in a single container?

Vince: The tenant is a concept for Okapi and folio. But Okapi and folio don't know docker. You have the ability to manage your Docker containers. You can load balance in two ways. You can put multiple modules and register multiple containers for the same module in Okapi. And you can choose which one you want to register your tenants into. Okapi will do the routing to the appropriate module, which is a docker itself or not. The other way is you can put a balancer in front of your modules. And you register the load balancer to Okapi. Okapi doesn't know the multiple modules. Okapi just knows there is one module. But in reality the module itself is not a module. It's a load balancer. And the load balancer distributes the load to the different modules you have. The second is the one we use at EBSCO. We use the features and capabilities of ECS to scale horizontally. If we have too much traffic, we add modules behind that balancer. And if we have less traffic, we can reduce the number behind. Okapi doesn't know the difference.

- Vince: There's no recommendations for managing Postgres. The database itself is a little bit abstracted through RMB. So the developers don't really have very much interaction with the database directly. It is going through the clients that are created

for the queries and everything else. When we go into deployment into AWS, typically we will use the AWS tools that are associated with the database being offered. So you can run some AWS tools to access and manage your database.

3. How does SHL solve the problem of bulk loading of user records into Folio?

- Gang Zhou: Records are loaded directly into the user database. The user data was converted to csv format, and copied into the database. It took around 2 hours for 5M+ user records.