Theme: Shanghai library FOLIO project

Time: July 28, 2020 07:00pm (EST) / July 29, 2020 07:00am (GMT+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. **Deploying modules with Docker**
   Gang Zhou: We are interested in deploying different modules by different Docker, but didn't find a docker repository on folio. Any plan for the deployment of docker by different modules?
   Vince: No current work or plan. Each institution has its own mechanism and works on its own. The most detailed description about the implementation of Docker and Rancher on wiki so far is by Jason Root from Texas A&M (https://wiki.folio.org/display/SYSOPS/Folio+on+Kubernetes+with+Rancher+2.1) (Notes: The link was added by Lucy. Vince, please let me know if this is not the link you referred to in the meeting).
   Sha Jiang: Should we put each module in one pod instead of combining them?
   Vince: How people deploy folio is up to each individual institution. Folio started simply. Initially there was no concern about clusters, pods, Kubernetes, etc. Instead, it cared about onsite deployment through Okapi and Java jar files. As the scale expands, it makes sense to put similar things into one pod (not one pod per module) to help with scaling. Modules that work together in the same domain are put in one pod, for example, mod-inventory and mod-inventory-storage in the same pod.
   Sha Jiang: Any community efforts on this?
   Vince: Texas A&M is doing Docker, rancher and Kubernetes deployment. EBSCO uses Docker. EBSCO also uses ECS from Amazon, but will move to Kubernetes. Modules are currently in the same cluster at EBSCO. We put different clusters for databases and different customers for Okapi and different modules. This meets our needs as a hosting provider for now. In the case that Shanghai is pursuing a massively scalable system, it makes sense to do what you suggested.

2. **Deploying a cluster for performance test at Shanghai Library**

Sha Jiang: We want to deploy modules in different pods. Any experience to share?

Vince: As far as I know, no one has gone beyond what Texas A&M has to offer. Good that you will be a pioneer.

Gang Zhou: If we have any questions about the cluster environment, can experts from EBSCO or the community help?

Vince: definitely.

Sha Jiang: Should Okapi be in a separate pod, or in each pod together with the modules?

Vince: Dedicate a cluster for Okapi alone, i.e. make one pod just for Okapi and run Okapi in cluster mode inside the pod. That will be used as the Okapi for all the other modules in the system in other pods.

Sha Jiang: We can't use Docker when we deploy modules in Okapi because Okapi can't run Docker in its own pod. Should we use openURL in the deployment descriptor?

Vince: No, because you use a deployment descriptor when you want Okapi to do the deployment for you. But we don't ask Okapi to do deployment.

Sha Jiang: How to deploy then?

Vince: You stand up your pods and your modules, and they already run. All you need to do is to register with Okapi.

Sha Jiang: Can Okapi control the lifecycle of modules?

Vince: No. In production it's not a good idea. Let the Rancher do the management.

3. Following up on joining the coreFunctional team

Lucy: Cate welcomed the Shanghai team to join coreFunctional. She shared a list of performance relevant issues in the backlog and asked the Shanghai team to evaluate. Tasks can't be prioritized based on Shanghai's selection. The Shanghai team may also propose new issues.

Sha Jiang: We want to know the workload and schedule.

Vince: It's better to discuss issues in a meeting to identify and prioritize tasks. Any arrangements for meetings?

Lucy: Not mentioned. Will ask Cate in tomorrow's meeting. - Done. Cate will post the regular coreFunctional meeting schedules (most meetings are scheduled 9pm or 11pm at Shanghai time) to the Slack channel #corefunctional_performance just in case Shanghai will be interested in attending them. Performance issues are not often discussed in the coreFunctional meetings. To save Shanghai team's time, we will share thoughts on the channel first. Then special meetings focused on performance can be arranged as needed.

4. ElasticSearch

Lucy: What do you think about the discussions between Mikhail and Sha Jiang on the Slack #searchengine channel?

Vince: Shanghai's plan is to use Logstash to make an advanced event, using Kafka to capture the events and using kafka messaging as a communication device

between Logstash, pushing data into ES for indexing. Mikhail didn't think Logstash was necessary. He advocated to keep using kafka as a message queue, but make direct connections to user, inventory, etc., push directly to the message queue and write a component for ES to retrieve from the message queue. Vince's comments: Both points are valuable. Another option here is to consider the Folio PubSub capabilities (mod-pubsub). It fits somewhat in the middle between the proposed approaches. It does not require the addition of kafka clients nor direct integration to Kafka for any module. It would still require the modification of individual modules to add calls to mod-pubsub for registration, publishing and subscribing. But this is a general purpose benefit and modules would ideally provide publications of their key events as a matter of course in Folio.

Sha Jiang: We will consider mod-pubsub.

To do: Vince will comment on the Slack channel #searchengine. Sha Jiang will change the proposed user stories accordingly and repost to the channel. Lucy will update information on JIRA.