Architecture

(C)Hurence

Version v1.0, 18.02.2020: First draft

Table des matières

Cette section décrit l’architecture du projet historian.

Descriptions des différents modules

Voici les différents modules de ce projet et leur rôle (voir les sections ci-dessous pour plus de détails)

- **gateway :** La gateway est notre serveur rest qui met à disposition notre API rest.
- **grafana-historian-datasource :** C'est un sous module git qui est notre datasource plugin grafana que nous utiliserons pour faire de la visualisation sur les timeseries de l'historian. La datasource sert de jonction entre grafana et la gateway.
- **historian-modele :** Contient les objets en commun entre les différents modules.
- **integration-tests :** Contient des classes utilitaires pour faire des tests d'intégrations.
- **loader :** Contient du code qui permet de recompacter des chunks déjà contenu dans l'historian.
- **logisland-timeseries :** Ce module étend le module "timeseries" avec des spécificités logisland.
- **timeseries :** Une library qui permet de manipuler des timeseries et de les compresser.

gateway

La gateway est implémenté avec vertx. Il propose une API rest que le grafana-historian-datasource utilise pour tracer les graphes dans grafana. Cela implique que ces deux modules sont étroitement liés, si on modifie l'api de la gateway cela aura un impacte sur la datazsource et vice versa.

grafana-historian-datasource

La datasource est implémenter en utilisant javascript, le projet github correspondant est disponible [ici](https://github.com/Hurence/grafana-historian-datasource), nous recommendons de développer ce plugin directement dans son propre projet. Une fois que vous avez push vos modifications dans le projet de la datasource, vous pouvez mettre à jour le sous module dans ce projet avec la commande suivante :

```
git submodule update --remote
```

Cela va automatiquement mettre à jour le projet avec le dernier commit du projet du sous module.

historian-modele

C'est un module qui sert à partager des class qui est utilisé par la gateway mais aussi d'autres modules. Par exemple le module loader. Cela évite de devoir embarquer la dépendance sur la gateway (qui contient beaucoup de choses).

integration-tests

Ce module contient des utilitaires pour fair des tests d'intégrations. Il contient notamment des extensions junit5 pour Spark et pour Solr. L'extension pour solr utilise les libraries

[testcontainers](https://www.testcontainers.org/), pour constituer un cluster solr avec docker compose. Ce module contient également dans ses ressources les configuration nécessaire pour créer les différentes collections utilisé par notre historian.

loader

Le loader contient des jobs batch. Pour l'instant il est implémenter un unique job qui sert a recompacter les chunks de l'historian avec différente tailles paramétrable. Par exemple lors de l'injection temps réel dans le data historian, on injecte des chunks de petites tailles, or pour optimiser les performances nous avons besoin de chunks de grosse tailles. De plus selon le type de requête effectuer et la fréquence des points pour une métrique données jouer sur la taille des chunks est très important.

Par exemple si l'on fait des requêtes sur de grande périodes, nous avons besoin de gros chunks pour gagner en performance, en effet nous n'aurons même pas besoin de décompresser ces chunks, utiliser leur agrégation précalculer sera suffisant. D'ailleurs ce job de recompaction pourra également servir de base pour nous permettre dans le futur de rajouter de nouvelles agrégations pré-calculer si nécessaire, ou bien de rajouter n'importe quel information qui nous paraît nécessaire.

Le compacteur utilise spark et solr.

logisland-timeseries

Contient des processeurs logisland utilisant la library timeseries. Ces processeurs manipulent des timeseries. Certains sont utilisés pour compacter des données dans le compacteur. A terme, il serait bien de supprimer ce module et de n'utiliser que des outils présent dans le module timeseries. En effet il ne paraît pas nécessaire d'embarquer des logiques métier logisland dans l'historian.

timeseries

C'est une library qui permet de manipuler les metrics d'une manière général. On peut par exemple les compresser et calculer des pré agrégations.