

BiJoe - Development #20762

Pouvoir utiliser des décomptes pré-calculer

18 décembre 2017 11:20 - Benjamin Dauvergne

Statut:	Nouveau	Début:	18 décembre 2017
Priorité:	Normal	Echéance:	
Assigné à:		% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:		Planning:	Non
Patch proposed:	Non		

Description

Actuellement on sait juste compter des choses dans une table via SELECT COUNT(id) FROM table [GROUP BY somegrouping] mais il pourra arriver qu'au lieu de cela on ait une colonne count ou le décompte aura déjà été fait, alors déjà il va falloir définir comment on regroupe les décomptes pour une façon de grouper, est-ce que ça a un sens ? Exemple de données.

timestamp	user_count	ou
2017-01-12	1345	usagers
2017-01-12	10	agent
2017-01-13	1347	usagers
2017-01-13	9	agent

- Comment je calcule le nombre d'usagers grouper par OU en ignorant le temps (je dirai en prenant la mesure la plus récente, i.e. ORDER BY timestamp LIMIT 1) ?
- Comment je calcule le nombre d'usager par mois (je dirai en gardant la dernière mesure du mois) ? [1]
- Comment je calcule la variation par jour (en utilisant la fonction lag() de Postgres [2]) ?

etc...

Et il faudra voir comment exprimer tout ça dans le format de configuration/schémas de bijoe.

<sup>1</sup><https://stackoverflow.com/questions/17436868/get-last-record-from-each-month>

<sup>2</sup><https://www.periscopedata.com/blog/computing-day-over-day-changes-with-window-functions>

Demandes liées:		
Lié à Authentic 2 - Development #20698: Fournir les statistiques authentic en...	Rejeté	14 décembre 2017

Historique

#1 - 18 décembre 2017 11:21 - Benjamin Dauvergne

- Description mis à jour

#2 - 18 décembre 2017 11:23 - Benjamin Dauvergne

- Description mis à jour

#3 - 18 décembre 2017 11:23 - Benjamin Dauvergne

- Lié à Development #20698: Fournir les statistiques authentic en olap pour exploitation via bi-joe. ajouté

#4 - 14 mars 2018 09:25 - Paul Marillonnet

- Assigné à mis à Paul Marillonnet

#5 - 14 mars 2018 11:30 - Paul Marillonnet

Benjamin Dauvergne a écrit :

alors déjà il va falloir définir comment on regroupe les décomptes pour une façon de grouper, est-ce que ça a un sens ?

Je ne comprends pas encore très bien quel niveau de généricité doit être supporté.  
Même interrogation pour le niveau de configuration, par l'utilisateur, des statistiques de décomptes pré-calculées produites.

**#6 - 14 mars 2018 11:42 - Paul Marillonnet**

Benjamin Dauvergne a écrit :

- Comment je calcule la variation par jour (en utilisant la fonction lag() de Postgres [2]) ?

Pour ce cas précis d'usage de la fonction lag(), en quoi a-t-on besoin de pré-calculer des décomptes ? (puisque visiblement on peut tout avoir en une requête, apparemment pas trop gourmande).

**#7 - 14 mars 2018 11:46 - Benjamin Dauvergne**

Je ne sais pas de quoi tu parles avec ce pré-calcul. La situation de base ici c'est qu'un logiciel ne nous transmet pas la liste des événements mais un état déjà agrégé (authentific fait cela). Il n'y a pas à déterminer si on doit ou non pré-calculer quelque chose.

**#8 - 14 mars 2018 11:54 - Paul Marillonnet**

Je ne comprends pas ce qu'il y a à déterminer dans ce cas.

Je pensais qu'il s'agissait de disposer d'une procédure de maintien d'une table de fluctuations temporelles pour les données agrégées reçues. C'est ça que je voyais comme le "pré-calcul" (au sens où ce serait calcul avant traitement final et affichage des données).

**#9 - 14 mars 2018 12:04 - Benjamin Dauvergne**

Non lag() s'utilise en temps réel (à chaque visu on refait un select avec lag()), on ne stocke pas le résultat dans une vue matérialisée ou une table temporaire).

**#10 - 14 mars 2018 12:06 - Benjamin Dauvergne**

Il y a à déterminer les requêtes SQL à lancer, bijoe pour l'instant ne sait générer que des requêtes SQL qui calculent des agrégats pas d'extraire des agrégats existants (c'est dans la description du ticket).

**#11 - 14 mars 2018 12:45 - Paul Marillonnet**

Donc, dans les grandes lignes :

- déterminer les requêtes SQL permettant d'avoir les résultats demandés dans la description du ticket
- évaluer la possibilité de générer ces requêtes
- déterminer la façon dont la configuration JSON de Bijoe pourrait être adaptée pour prendre en charge le paramétrage de ces requêtes génériques

C'est bien de ça qu'il s'agit ?

**#12 - 14 mars 2018 13:51 - Benjamin Dauvergne**

Oui.

**#13 - 09 août 2018 12:38 - Paul Marillonnet**

Donc, si j'ai un ensemble de requêtes du genre :

```
select count(custom_user_user.uuid), a2_rbac_organizationalunit.name from custom_user_user left join a2_rbac_organizationalunit on custom_user_user.ou_id = a2_rbac_organizationalunit.id group by a2_rbac_organizationalunit.id;
```

il faut trouver le moyen de les exécuter périodiquement pour le remplissage de la table de pré-calcul, c'est bien ça ?

**#14 - 09 août 2018 13:49 - Benjamin Dauvergne**

Non, on ne s'occupe pas ici de comment on obtient les données, c'est comment bijoe pourrait les exploiter qui nous intéresse.

**#15 - 09 août 2018 17:41 - Paul Marillonnet**

D'accord, je croyais nécessaire de se mettre d'abord d'accord sur les requêtes qui arrivent en entrée.

Donc ici ce serait surtout du travail sur bijoe.schemas et bijoe.engine pour faire en sorte qu'on puisse traiter de telles données en entrée ? (Ensuite, si le code est bon, on crée les bijoe.engine.Cube et le bijoe.schemas.Warehouse qui conviennent)

Je me plante ?

**#16 - 21 septembre 2020 16:53 - Pierre Cros**

- Statut changé de Nouveau à Information nécessaire

- Assigné à changé de Paul Marillonnet à Benjamin Dauvergne

Benj, je te file ce ticket pour que tu orientes Paul.

**#17 - 08 octobre 2020 13:43 - Benjamin Dauvergne**

- *Statut changé de Information nécessaire à Nouveau*

- *Assigné à Benjamin Dauvergne supprimé*

On abandonne bijoe.