

w.c.s. - Development #9553

requêtes SQL pour les graphes des stats globales

08 janvier 2016 15:30 - Frédéric Péters

Statut:	Fermé	Début:	08 janvier 2016
Priorité:	Bas	Echéance:	
Assigné à:		% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:	v1.36	Planning:	
Patch proposed:	Oui		
Description			
duration: 4450.041 ms statement: SELECT DATE_TRUNC('month', receipt_time) AS month, COUNT(*) FROM wcs_all_forms WHERE receipt_time IS NOT NULL GROUP BY month ORDER BY month			
duration: 4651.421 ms statement: SELECT DATE_TRUNC('year', receipt_time) AS year, COUNT(*) FROM wcs_all_forms WHERE receipt_time IS NOT NULL GROUP BY year ORDER BY year			
duration: 4481.625 ms statement: SELECT DATE_PART('dow', receipt_time) AS weekday, COUNT(*) FROM wcs_all_forms WHERE receipt_time IS NOT NULL GROUP BY weekday ORDER BY weekday			
duration: 4569.341 ms statement: SELECT DATE_PART('hour', receipt_time) AS hour, COUNT(*) FROM wcs_all_forms WHERE receipt_time IS NOT NULL GROUP BY hour ORDER BY hour			
De ce que j'ai pu lire, la bonne manière pour optimiser ça c'est d'utiliser des "Materialized Views" (http://www.postgresql.org/docs/9.4/static/rules-materializedviews.html), sur lesquelles il sera possible de poser des index sur les différents morceaux de la date.			
Avec sur le côté, de manière régulière un REFRESH MATERIALIZED VIEW CONCURRENTLY (http://www.postgresql.org/docs/9.4/static/sql-refreshmaterializedview.html).			
Ça devrait sans doute exister en parallèle de l'actuel wcs_all_forms car on a besoin de données à jour pour la vue globale.			
(ça doit de toute façon attendre qu'on puisse utiliser postgresql 9.4)			
Demandes liées:			
Dupliqué par w.c.s. - Bug #10174: longues requetes SQL → 504 Gateway Time-out		Rejeté	03 mars 2016

Révisions associées

Révision 98ae3b05 - 04 mars 2016 15:27 - Frédéric Péters

sql: use materialized views if available, for statistics (#9553)

Révision 2e130848 - 04 mars 2016 16:00 - Frédéric Péters

sql: don't use 'IF NOT EXISTS' when creating materialized view (#9553)

(it requires postgresql 9.5)

Historique

#1 - 08 janvier 2016 15:55 - Benjamin Dauvergne

À noter aussi qu'avec 9.5 on pourra en plus calculer toute sorte d'aggrégats en une seule requête¹.

¹<http://www.depesz.com/2015/05/24/waiting-for-9-5-support-grouping-sets-cube-and-rollup/>

#2 - 16 janvier 2016 13:27 - Frédéric Péters

J'ai tenté différentes choses alternatives, quand même :

- créer un index sur le receipt_time
- créer un index sur le date_trunc('xxx', receipt_time) sur chacune des tables
- créer une colonne receipt_month qui soit le date_trunc(...) sur chacune des vues

Mais non, postgresql ne trouve pas comment utiliser ça pour optimiser sa requête; toujours il assemble d'abord toutes les lignes de toutes les tables.

#3 - 16 janvier 2016 23:04 - Benjamin Dauvergne

Même en créant l'index sur la vue matérialisée ?

#4 - 17 janvier 2016 12:04 - Frédéric Péters

Il n'y a pas encore de vue matérialisée, nos serveurs sont encore en 9.1.

#6 - 03 mars 2016 14:47 - Frédéric Péters

- Dupliqué par Bug #10174: longues requetes SQL → 504 Gateway Time-out ajouté

#7 - 03 mars 2016 14:57 - Thomas Noël

à noter que postgresql 9.4 est dans wheezy-backports

#8 - 03 mars 2016 18:34 - Frédéric Péters

- Fichier 0001-sql-use-materialized-views-if-available-for-statisti.patch ajouté

- Statut changé de Nouveau à En cours

- Patch proposed changé de Non à Oui

Patch qui regarde au runtime la version de postgresql. Testé en local sur les données de Vincennes, il n'y a pas besoin de fixer d'index pour que ça soit tout de suite très performant.

#9 - 04 mars 2016 15:05 - Thomas Noël

Ack

#10 - 04 mars 2016 15:27 - Frédéric Péters

- Statut changé de En cours à Résolu (à déployer)

```
commit 98ae3b0581dea73a171a7f81e671b45a7e963f35
Author: Frédéric Péters <fpeters@entrouvert.com>
Date: Thu Mar 3 18:33:07 2016 +0100
```

```
sql: use materialized views if available, for statistics (#9553)
```

#11 - 04 mars 2016 15:49 - Frédéric Péters

- Fichier 0001-sql-don-t-use-IF-NOT-EXISTS-when-creating-materializ.patch ajouté

#12 - 04 mars 2016 15:56 - Thomas Noël

J'aurais bien vu un "if pg_version >= 95000: ajouter if not exists"

#13 - 04 mars 2016 15:57 - Thomas Noël

mais sinon, ack

#14 - 04 mars 2016 16:01 - Frédéric Péters

```
commit 2e130848b7c2efe21d38085d2592566b4ceabf87
Author: Frédéric Péters <fpeters@entrouvert.com>
Date: Fri Mar 4 15:48:04 2016 +0100
```

```
sql: don't use 'IF NOT EXISTS' when creating materialized view (#9553)
```

```
(it requires postgresql 9.5)
```

#15 - 04 mars 2016 18:04 - Thomas Noël

- Version cible mis à v1.36

#16 - 04 avril 2016 10:10 - Thomas Noël

- Statut changé de Résolu (à déployer) à Fermé

Fichiers

0001-sql-use-materialized-views-if-available-for-statisti.patch	9,29 ko	03 mars 2016	Frédéric Péters
0001-sql-don-t-use-IF-NOT-EXISTS-when-creating-materializ.patch	1,01 ko	04 mars 2016	Frédéric Péters