OLAP / Business Intelligence pour Publik - Development #19422

temps pris pour récupérer les demandes

13 octobre 2017 14:09 - Serghei Mihai

Statut:	Nouveau	**Début:**	13 octobre 2017
Priorité:	Normal	**Echéance:**	
Assigné à:		**% réalisé:**	0%
Catégorie:		**Temps estimé:**	0:00 heure
Version cible:			
Patch proposed:	Non	**Planning:**	

Description

Avec le nombre croissant d'instances de bijoe deployés et des demandes "lourdes" dans wcs il faudrait réduire la fréquence d'interrogation de wcs: passer d'une heure à genre 3 heures.

Demandes liées:

Lié à Publik - Development #22442: durée des wcs-olap	**Rejeté**	**11 mars 2018**
Lié à OLAP / Business Intelligence pour Publik - Development #56039: améliore...	**Fermé**	**07 août 2021**
Dupliqué par OLAP / Business Intelligence pour Publik - Bug #36669: optimiser...	**Rejeté**	**03 octobre 2019**

Historique

#1 - 13 octobre 2017 14:23 - Frédéric Péters

On ne peut pas commencer par multiplier par trois les performances ?

#2 - 13 octobre 2017 15:11 - Frédéric Péters

Notamment je lis dans #18980 "2348 demandes avec toutes les pièces jointes dedans" et je ne pense pas qu'il y ait d'utilité à récupérer la moindre pièce jointe.

#3 - 13 octobre 2017 15:37 - Benjamin Dauvergne

Il y a ça dans la méthode json() de wcs.backoffice.management.FormPage :

```
            output = [filled.get_json_export_dict(include_files=False, anonymise=anonymise)
                      for filled in items]
```

je ne pense pas qu'on récupère les pièces jointes, ou bien tu soulignes juste le commentaire mal avisé de Serghei ?

#4 - 13 octobre 2017 16:16 - Frédéric Péters

Je pointais juste le commentaire sans regarder le code.

#5 - 13 octobre 2017 16:20 - Serghei Mihai

Ce ticket n'a pas de rapport avec le cas d'Alfortville.
Ici, je propose juste de réduire la fréquence de récuperation des demandes en vue du nombre croissant et on peut tout à fait partir sur l'amélioration des performances.

Je ne pense pas non plus que les pièces jointes sont récuperées, mais que la récuperation du formdata contenant des fichiers est plus longue et donc wcs-olap se prend un timeout (http 502), d'ou la réduction du batch_size.

#6 - 16 octobre 2017 12:07 - Serghei Mihai

Dans l'idée de multiplier les performances on devrait peut-être aussi regarder dans wcs:

```
5.135.221.21 - - [16/Oct/2017:11:15:29 +0200] "GET /api/forms/demande-de-place-en-creche/list?anonymise&full=o
n&offset=1100&limit=100&orig=statistiques.alfortville.fr&algo=sha256&timestamp=2017-10-16T09%3A15%3A12Z&nonce=
cff7e8e058d9a206730061113a630505&signature=nNBK26V1sli8OJSF8ocLqv4L6v79jVMBdIqcEO3ghm4%3D HTTP/1.1" 200 176993
53 "-" "python-requests/2.4.3 CPython/2.7.9 Linux/2.6.32-48-pve" "demarches.alfortville.fr" [17.357 s]
5.135.221.21 - - [16/Oct/2017:11:15:46 +0200] "GET /api/forms/demande-de-place-en-creche/list?anonymise&full=o
n&offset=300&limit=100&orig=statistiques.alfortville.fr&algo=sha256&timestamp=2017-10-16T09%3A15%3A39Z&nonce=e
84e552fd44671507f23fd2c0420d8b7&signature=LdXf%2Bldd8Q%2BLPA3HgSA0svXpOBzrSYc5NLjSFO1XwRs%3D HTTP/1.1" 200 490
2599 "-" "python-requests/2.4.3 CPython/2.7.9 Linux/2.6.32-48-pve" "demarches.alfortville.fr" [6.347 s]
5.135.221.21 - - [16/Oct/2017:11:20:04 +0200] "GET /api/forms/demande-de-place-en-creche/list?anonymise&full=o
n&offset=1400&limit=100&orig=statistiques.alfortville.fr&algo=sha256&timestamp=2017-10-16T09%3A19%3A52Z&nonce=
```

```
c88a0558f7f504302c76c6e161b2fce8&signature=HJqt9709t7uKTw%2BixkRv5E0kytAOKahPBgcu%2BsirN54%3D HTTP/1.1" 200 12
326595 "-" "python-requests/2.4.3 CPython/2.7.9 Linux/2.6.32-48-pve" "demarches.alfortville.fr" [11.706 s]
5.135.221.21 - - [16/Oct/2017:11:20:28 +0200] "GET /api/forms/demande-de-place-en-creche/list?anonymise&full=o
n&offset=1000&limit=100&orig=statistiques.alfortville.fr&algo=sha256&timestamp=2017-10-16T09%3A20%3A13Z&nonce=
23fddd0234b502bb5f0bcd5a70a04700&signature=bXlxSdVt9wOgvLP2A7upsnEcRwyyBpxVp1Los4tJIp0%3D HTTP/1.1" 200 159472
06 "-" "python-requests/2.4.3 CPython/2.7.9 Linux/2.6.32-48-pve" "demarches.alfortville.fr" [15.276 s]
```

Récuperer 100 demandes en 17 secondes, ça me paraît beaucoup.

#7 - 16 octobre 2017 12:09 - Frédéric Péters

À noter les tailles des réponses, 17699353 pour la première.

#8 - 16 octobre 2017 12:58 - Benjamin Dauvergne

Il y un souci si la réponse fait 17Mo, 17s pour 100 demandes, ça fait 170Ko seconde c'est pas si mal, il faudrait savoir ce qu'il y a dans ces 170Ko.

#9 - 11 mars 2018 13:49 - Frédéric Péters

- Sujet changé de réduire la fréquence de récuperation des demandes à temps pris pour récupérer les demandes

Je change l'intitulé parce que l'objectif ne doit pas être de réduire la fréquence et au vu des commentaires précédents il y a des trucs à explorer.

#10 - 11 mars 2018 13:50 - Frédéric Péters

- Lié à Development #22442: durée des wcs-olap ajouté

#11 - 11 mars 2018 14:16 - Frédéric Péters

> Il y un souci si la réponse fait 17Mo, 17s pour 100 demandes, ça fait 170Ko seconde c'est pas si mal, il faudrait savoir ce qu'il y a dans ces 170Ko.

J'ai donc rejoué ces requêtes.

L'espace consommé, c'est des historiques super longs, avec un statut bouclant sur lui-même (sans doute sur une vérification de webservice).

ex:

```
        {
           "time" : "2016-11-10T14:02:23Z",
           "status" : "finished"
        },
        {
           "status" : "finished",
           "time" : "2016-11-10T15:03:10Z"
        },
        {
           "time" : "2016-11-10T16:23:00Z",
           "status" : "finished"
        },
        {
           "status" : "finished",
           "time" : "2016-11-10T17:43:19Z"
        },
        {
           "time" : "2016-11-10T18:43:20Z",
           "status" : "finished"
        },
```

Ça rejoint #22236.

Mais côté performances pures, ça pourrait aussi être amélioré par #22234.

...

Bon, **pourrait**, mais en fait non comme on en charge un tas le gain apparait insignifiant. (après avoir posé un index manuellement sur la démarche testée)

#12 - 20 mars 2018 11:35 - Frédéric Péters

En attendant, la mort dans l'âme,

```
-0 * * * * bijoe /usr/lib/bijoe/import-wcs-data.sh
+0 3 * * * bijoe /usr/lib/bijoe/import-wcs-data.sh
```

#13 - 27 novembre 2019 18:40 - Frédéric Péters

- Dupliqué par Bug #36669: optimiser le temps de contruction des données ajouté

#15 - 07 août 2021 19:15 - Benjamin Dauvergne

- Lié à Development #56039: améliorer les performances du chargement ajouté