# Hobo - Development #67377

## moteurs d'indexation qui tournent en rond avec un paramètre ?next de plus en plus long

14 juillet 2022 20:40 - Frédéric Péters

| | | | |
|---|---|---|---|
| **Statut:** | Fermé | **Début:** | 14 juillet 2022 |
| **Priorité:** | Normal | **Echéance:** | |
| **Assigné à:** | Benjamin Dauvergne | **% réalisé:** | 0% |
| **Catégorie:** | | **Temps estimé:** | 0:00 heure |
| **Version cible:** | | | |
| **Patch proposed:** | Oui | **Planning:** | Non |

**Description**

```
54.36.149.40 - - [14/Jul/2022:20:24:13 +0200] "GET /login/?next=%2Fpassword%2Freset%2F%3Fnext%3D
%2Flogin%2F%253Fnext%253D%25252Flogin%25252F%25253Fnext%25253D%2525252Faccounts%2525252Fpassword
%2525252Freset%2525252F%2525253Fnext%2525253D%2525252Flogin%2525252F%252525253Fnext%252525253D%2
5252525252Flogin%25252525252F%252525252553Fnext%25252525253D%2525252525252Flogin%2525252525252F%2
525252525253Fnext%2525252525253D%252525252525252Faccounts%252525252525252Fpassword%2525252525252
52Freset%252525252525252F%252525252525253Fnext%252525252525253D%252525252525252Flogin%2525252525
25252F%2525252525252525253Fnext%25252525252525253D%252525252525252525252Flogin%252525252525252525252F%
252525252525252525253Fnext%25252525252525252525253D%252525252525252525252525252Flogin%252525252525252525252F
%252525252525252525253Fnext%25252525252525252525253D%252525252525252525252525252Flogin%252525252525252525252
52525252F%252525252525252525252525253Fnext%252525252525252525252525253D%252525252525252525252525252Flogin%
252525252525252525252525252F%252525252525252525252525252525253Fnext%252525252525252525252525253D%25252525252
5252525252525252Flogin%252525252525252525252525252F%252525252525252525252525252525253Fnext%25252525252
5252525252525253D%252525252525252525252525252525252Faccounts%252525252525252525252525252525252Fpassword%
252525252525252525252525252Freset%252525252525252525252525252F%252525252525252525252525252525253
Fnext%252525252525252525252525252525253D%252525252525252525252525252525252Faccounts%252525252525252525
252525252Fpassword%252525252525252525252525252525252Freset%252525252525252525252525252F%2525252525
252525252525252525253Fnext%252525252525252525252525252525253D%252525252525252525252525252525252Fidp%25
252525252525252525252525252Fsaml2%252525252525252525252525252525252Fcontinue%252525252525252525252525
5252525253Fnonce%252525252525252525252525252525253D_0CA9192F060E1B9AE5AAEB8AA21BC523 HTTP/1.1" 200
 4095 "-" "Mozilla/5.0 (compatible; AhrefsBot/7.0; +http://ahrefs.com/robot/)"
```

Bien sûr c'est le moteur qui est un peu bête mais il y aurait peut-être quelque chose à faire pour éviter ça ?

Dans l'access.log du jour, 75% des lignes contiennent %2525252525 (25 étant % quoté, la répétition étant que % s'est trouvé encodé en %25, qui s'est trouvé encodé en %2525, qui en %252525, etc.).

(85% de lignes avec 252525, 81 avec %25252525).

Un truc avec très peu d'intelligence, type middleware qui simplement retournera une 400 en présence de %252525, ça pourrait arrêter les moteurs plus tôt, et j'ai vérifié il n'y a pas de faux positif.

### Révisions associées

**Révision 48f1e920 - 09 août 2022 09:40 - Benjamin Dauvergne**

context_processors: keep next_url for user urls on authentic (#67377)

### Historique

**#1 - 14 juillet 2022 20:43 - Frédéric Péters**

*- Description mis à jour*

**#3 - 15 juillet 2022 10:26 - Benjamin Dauvergne**

*- Projet changé de Authentic 2 à Hobo*

*- Assigné à mis à Benjamin Dauvergne*

C'est du aux liens connexions/enregistrement dans le header qui reprennent bêtement le chemin de la page locale comme ?next= sans voir qu'il y a déjà un ?next= dans la page en cours, les bots suivent /login -> /password-reset -> login -> etc..

**#4 - 15 juillet 2022 10:30 - Benjamin Dauvergne**

*- Fichier 0001-context_processors-keep-next_url-for-user-urls-on-au.patch ajouté*

*- Statut changé de Nouveau à Solution proposée*

*- Patch proposed changé de Non à Oui*

#### #5 - 29 juillet 2022 11:09 - Benjamin Dauvergne

*- Fichier 0001-context_processors-keep-next_url-for-user-urls-on-au.patch ajouté*

authentic2.utils.misc.get_next_url n'est pas vraiment disponible dans la cible hobo, on va donc mocker ça.

#### #6 - 09 août 2022 09:39 - A. Berriot

*- Statut changé de Solution proposée à Solution validée*

Si j'ai bien compris on laisse authentic parser le next et on utilise celui-là s'il existe. Si authentic ne voit pas de paramètre next, on utilise le chemin actuel comme next (comportement existant). J'imagine que ça coupe le problème à la racine, puisque ça stoppe les concaténations successives, et si le moteur d'indexation est pas trop mal fichu, il va voir qu'il a déjà crawlé les urls en question et passer son chemin.

#### #7 - 09 août 2022 09:41 - Frédéric Péters

*- Statut changé de Solution validée à Résolu (à déployer)*

```
commit 48f1e920bc571383dda07ae5d40e651666b23e41
Author: Benjamin Dauvergne <bdauvergne@entrouvert.com>
Date:   Fri Jul 15 10:30:15 2022 +0200

    context_processors: keep next_url for user urls on authentic (#67377)
```

#### #8 - 09 août 2022 10:14 - Transition automatique

*- Statut changé de Résolu (à déployer) à Solution déployée*

#### #9 - 14 août 2022 12:42 - Frédéric Péters

Ça n'a pas d'effet sur les URL déjà indexées avec des paramètres longs, pas de baisse constatée du nombre de visites; ex:

```
185.191.171.24 - - [14/Aug/2022:12:32:07 +0200] "GET /login/?next=%2Flogin%2F%3Fnext%3D%2Faccounts%2Fregister%
2F%253Fnext%253Dhttps%25253A%25252F%25252Fconnexion-amiens.test.entrouvert.org%25252Faccounts%25252Fregister%2
5252F%25253Fnext%25253D%2525252Flogin%2525252F%2525253Fnext%2525253D%2525252Faccounts%2525252Fregister%2525252
F%252525253Fnext%252525253Dhttps%25252525253A%25252525252F%25252525252Fconnexion-amiens.test.entrouvert.org%25
252525252Faccounts%25252525252Fregister%25252525252F%25252525253Fnext%25252525253Dhttps%2525252525253A%2525252
525252F%2525252525252Fconnexion-amiens.test.entrouvert.org%2525252525252Faccounts%2525252525252Fregister%25252
52525252F%2525252525253Fnext%2525252525253Dhttps%252525252525253A%252525252525252F%252525252525252Fconnexion-a
miens.test.entrouvert.org%252525252525252Faccounts%252525252525252Fregister%252525252525252F%252525252525253Fn
ext%252525252525253Dhttps%2525252525252525253A%2525252525252525252F%2525252525252525252Fconnexion-amiens.test.entrou
vert.org%2525252525252525252Faccounts%2525252525252525252Fregister%2525252525252525252F%2525252525252525253Fnext%25252
525252525253Dhttps%252525252525252525253A%252525252525252525252F%252525252525252525252Fconnexion-amiens.test.entrouv
ert.org%252525252525252525252Faccounts%252525252525252525252Fregister%252525252525252525252F%252525252525252525253Fnex
t%252525252525252525253Dhttps%25252525252525252525253A%25252525252525252525252F%25252525252525252525252Fconnexion-amie
ns.test.entrouvert.org%25252525252525252525252Faccounts%25252525252525252525252Fregister%25252525252525252525252F%25
252525252525252525253Fnext%25252525252525252525253Dhttps%2525252525252525252525253A%252525252525252525252525252F%2525252
5252525252525252Fconnexion-amiens.test.entrouvert.org%25252525252525252525252Faccounts%25252525252525252525252
Fregister%25252525252525252525252F%252525252525252525252525253Fnext%252525252525252525252525253D%25252525252525252525252
5252Fidp%252525252525252525252525252Fsaml2%25252525252525252525252525252Fcontinue%2525252525252525252525252525253Fnonce%25
252525252525252525252525253D_95158C6CADEB47D79A29D402E8760BFF%252525252525252525252526service%25252525252525252525252
253Dporte-doc HTTP/1.1" 200 3785 "-" "Mozilla/5.0 (compatible; SemrushBot/7~bl; +http://www.semrush.com/bot.ht
ml)" "connexion-amiens.test.entrouvert.org" [0.114 s]
```

#### #11 - 16 octobre 2022 04:42 - Transition automatique

Automatic expiration

### Fichiers

| | | | |
|---|---|---|---|
| 0001-context_processors-keep-next_url-for-user-urls-on-au.patch | 971 octets | 15 juillet 2022 | Benjamin Dauvergne |
| 0001-context_processors-keep-next_url-for-user-urls-on-au.patch | 3,59 ko | 29 juillet 2022 | Benjamin Dauvergne |