

Hobo - Bug #32435

validate_email_address : ça marchera pas...

18 avril 2019 09:02 - Thomas Noël

Statut:	Fermé	Début:	18 avril 2019
Priorité:	Normal	Echéance:	
Assigné à:	Christophe Siraut	% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:		Planning:	Non
Patch proposed:	Oui		

Description

Sur le principe : on ne peut pas vérifier une adresse email, parce que les serveurs répondent souvent en greylist (400). Donc ce code, pour moi, doit être retiré.

Mais si vous insistez, deux pépins, d'abord lors de la recherche du mx :

```
mx_server = dns.resolver.query(email_domain, 'MX')[0].exchange.to_text()
```

... la première réponse n'est pas le MX de moindre coût, donc ça risque de planter souvent, soit parce que le secondaire n'est pas là, soit par faux positif (la plupart des mx secondaires acceptent n'importe quel rcpt, ils ne peuvent pas vérifier).

Ensuite, smtpplib.SMTP : le timeout est trop long, risque de dépasser celui de nginx/uwsgi qu'on met habituellement (30 secondes): il faut mettre 20 ou 25 secondes max (et 10 ça suffit bien largement, la personne devant son navigateur aura déjà tapé trois fois sur F5 ;)).

Révisions associées

Révision 4a5b60a9 - 18 avril 2019 14:28 - Christophe Siraut

emails/validators: miscellaneous adjustments (#32435)

- address preferred mx server * set connection timeout * do not raise validation error on temporary failures * add an option to bypass smtp validation * catch all dns exceptions

Historique

#1 - 18 avril 2019 09:17 - Christophe Siraut

Pour ce qui est de ta première objection, le rejet greylisting (4xx) est retourné après la commande data; ici nous quittons avant.

#2 - 18 avril 2019 09:22 - Thomas Noël

Christophe Siraut a écrit :

Pour ce qui est de ta première objection, le rejet greylisting (4xx) est retourné après la commande data; ici nous quittons avant.

Le greylist c'est aussi possible plus tôt ; exemple sur un postgrey:

```
220 smtp.fr.auf.org ESMTP Postfix (Debian/GNU AUF Paris)
HELO coucou
250 smtp.fr.auf.org
MAIL FROM: tnoel@entrouvert.com
250 2.1.0 Ok
RCPT TO: abuse@auf.org
450 4.2.0 <abuse@auf.org>: Recipient address rejected: Greylisted, see http://postgrey.schweikert.ch/help/auf.org.html
QUIT
```

#3 - 18 avril 2019 09:26 - Thomas Noël

J'ajoute une couche: ça marchera pas sur les machines qui n'ont pas accès à smtp (celles qui doivent passer par un mta local pour envoyer des mails, cas fréquent en hébergement sur site)

#4 - 18 avril 2019 09:50 - Christophe Siraut

Tant qu'à faire autant virer tout 1b9704742f. (Une autre option est de ne pas lever d'erreur au moment de la modification de l'adresse, et d'ajouter un

v/x indicatif à sa droite si l'adresse peut être validée, mais la valeur ajoutée de cette modif est faible)

#5 - 18 avril 2019 10:05 - Frédéric Péters

1b9704742f (qui est l'unique commit de [#24519](#), donc tout ce qui touche la vérification d'email) contient également la vérification SPF et cette vérification est pour moi importante.

À noter dans la description de [#24519](#) : "et si ça répond en 5xx" (avec oui une hésitation entre parenthèse derrière dans une suggestion d'aller plus loin).

Bref, pour moi, [#24519](#) reste valide, il faut juste se limiter à 5xx (vraie erreur définitive), et faire ça avec un timeout court, et gérer la présence de plusieurs MX, et pouvoir débrayer pour les situations d'hébergement sur site nous refusant un accès au port 25 d'un serveur extérieur.

#6 - 18 avril 2019 10:16 - Thomas Noël

- Assigné à mis à Thomas Noël

Yep yep. Je regarde tout ça.

#7 - 18 avril 2019 10:50 - Benjamin Dauvergne

Pour le greylisting, je dirai qu'on peut rien faire à part être permissif (mais voir le point suivant).

Donc on doit tester tous les MX autant que faire se peut, si un est ne répond pas tant pis, si un seul refuse, on refuse. En gros :

```
ok = True
for mx in mxs:
    test with mx:
        if greylisted :
            ok = None
        elif unavailable:
            ok = None
        elif mail not accepted:
            ok = False
            break
        elif mail accepted and ok is None:
            ok = True
```

Si None on peut toujours mettre une alerte (je ne sais pas trop comment avec la logique de validation) qu'on est pas bien sûr, si ok=True ça va, si ok=False on refuse fermement.

Pour les cas où on a pas accès au port 25 en sortie, faut juste un setting HOBO_VALIDATE_EMAIL_WITH_SMTP = False,

On pourrait aussi imposer l'envoi d'un mail de validation avec un jeton et ne garder que la validation de syntaxe et SPF, le mail ne sera accepté que si on le rentre accompagné du jeton.

#8 - 18 avril 2019 10:59 - Christophe Siraut

on doit tester tous les MX autant que faire se peu

Pourquoi n'est-ce pas correct de s'adresser au MX qui a la priorité la plus haute?

#9 - 18 avril 2019 11:04 - Thomas Noël

Christophe Siraut a écrit :

on doit tester tous les MX autant que faire se peu

Pourquoi n'est-ce pas correct de s'adresser au MX qui a la priorité la plus haute?

Si si, c'est ce qu'il faut faire. Les MX secondaires ne peuvent en général pas savoir si l'adresse existe réellement (ce sont souvent juste des spool d'attente du mx principal)

#10 - 18 avril 2019 11:04 - Thomas Noël

- Assigné à changé de Thomas Noël à Christophe Siraut

#11 - 18 avril 2019 11:10 - Christophe Siraut

- Fichier 0002-emails-validators-address-preferred-mx-server-32435.patch ajouté
- Fichier 0003-emails-validators-set-connection-timeout-32435.patch ajouté
- Fichier 0004-emails-validators-do-not-raise-validation-error-on-t.patch ajouté
- Fichier 0005-emails-validators-add-option-to-bypass-smtp-validati.patch ajouté
- Statut changé de Nouveau à Solution proposée
- Patch proposed changé de Non à Oui

Voici les correctifs:

- s'adresser au serveur MX principal
- timeout à 10s
- lever une erreur seulement en cas d'erreur 5xx
- ajouter une option de débrayage

Je squasherai tout ça une fois validés.

#12 - 18 avril 2019 11:46 - Benjamin Dauvergne

Christophe Siraut a écrit :

Voici les correctifs:

- s'adresser au serveur MX principal

Le principe des MX veut qu'on s'adresse à tous les MX, pas seulement au premier, non ?

En fait j'ai du mal à faire le lien avec ce fonctionnement et la remarque de Thomas comme quoi au milieu de tout ça on peut avoir de simples relais avec possiblement la même priorité que les autres et qui vont toujours dire OK, on est vraiment pas sûr que ce soit configuré comme nous on aimerait, je ne vois pas d'autre solution que de s'adresser à tout le monde et d'ignorer les 4xx/erreur de connection. Et en même temps je ne sais pas si on peut faire la différence entre un refus de relai permanent (parce que le SMTP n'est plus un MX pour ce domaine mais c'est mal configuré) et un refus de relai parce que l'email n'est pas bonne.

À la fin la bonne vieille vérité c'est que le seul moyen de savoir qu'une email existe c'est encore de lui envoyer un mail.

#13 - 18 avril 2019 12:05 - Thomas Noël

- 0002-emails-validators-address-preferred-mx-server-32435.patch : ok (bon si on était des supermen, le test ajouterait un mx bidon avec un coût de 20 en premier dans la liste, genre [mx20, mx10])
- 0003-emails-validators-set-connection-timeout-32435.patch : ok
- 0004-emails-validators-do-not-raise-validation-error-on-t.patch : il faut poser le smtp.quit() juste après le smtp.rcpt(value), puisqu'on le fait de toute façon. Pour la condition "status >= 500" on écrit plus souvent "status // 100 == 5", mais après tout >=500 ça va aussi attraper les erreurs 600 et 700 alors c'est bon ;) (je ne retrouve plus la blague sur les faux status genre 747 = trop gros message, etc)
- 0005-emails-validators-add-option-to-bypass-smtp-validati.patch : ok (on pourrait aussi avoir un autre flag qui permette de bypasser le check SPF sur les hébergements où les requêtes DNS hors zones locales ne marchent pas (c'était le cas au cd14 du temps de l'hébergement sur site... mais bon, on peut attendre que ça revienne ailleurs, jamais revu ça))
- et aussi ajouter aussi un patch "except dns.exception.DNSException as e:" lors de l'appel à dns.resolve, qui remplacera les except sur NXDOMAIN et NoAnswer (dans le code actuel par exemple on gère pas dns.resolve.NoNameservers = domaine qui n'existe pas... autant catcher toutes les exception et afficher le message et zou).

#14 - 18 avril 2019 12:06 - Thomas Noël

Benjamin Dauvergne a écrit :

Le principe des MX veut qu'on s'adresse à tous les MX, pas seulement au premier, non ?

Nope, un par un, dans l'ordre de coût (du moins cher au plus cher). Ici on s'adresse au premier, si ça foire on laisse béton, c'est normalement un cas très rare alors j'accepte qu'il se produise (sinon on peut ré-écrire un client mail, mais merci)

#15 - 18 avril 2019 12:39 - Benjamin Dauvergne

Thomas Noël a écrit :

Benjamin Dauvergne a écrit :

Le principe des MX veut qu'on s'adresse à tous les MX, pas seulement au premier, non ?

Nope, un par un, dans l'ordre de coût (du moins cher au plus cher). Ici on s'adresse au premier, si ça foire on laisse béton, c'est normalement un

cas très rare alors j'accepte qu'il se produise (sinon on peut ré-écrire un client mail, mais merci)

Ok.

#16 - 18 avril 2019 12:41 - Christophe Siraut

- Fichier 0002-emails-validators-address-preferred-mx-server-32435.patch ajouté
- Fichier 0003-emails-validators-set-connection-timeout-32435.patch ajouté
- Fichier 0004-emails-validators-do-not-raise-validation-error-on-t.patch ajouté
- Fichier 0005-emails-validators-add-option-to-bypass-smtp-validati.patch ajouté
- Fichier 0006-emails-validators-catch-all-dns-exceptions-32435.patch ajouté

Avec les adaptations de Thomas.

C'est mieux de tout squasher avant de pousser?

#17 - 18 avril 2019 14:02 - Thomas Noël

- Statut changé de Solution proposée à Solution validée

Christophe Siraut a écrit :

Avec les adaptations de Thomas.

Ack.

C'est mieux de tout squasher avant de pousser?

Comme tu veux :)

#18 - 18 avril 2019 14:30 - Christophe Siraut

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

```
commit 4a5b60a95a6b90745e4b674ea2a58bbf90c938ff
Author: Christophe Siraut <csiraut@entrouvert.com>
Date: Thu Apr 18 09:33:34 2019 +0200
```

```
emails/validators: miscellaneous adjustments (#32435)
```

- * address preferred mx server
- * set connection timeout
- * do not raise validation error on temporary failures
- * add an option to bypass smtp validation
- * catch all dns exceptions

#19 - 19 avril 2019 11:15 - Frédéric Péters

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

Fichiers

0002-emails-validators-address-preferred-mx-server-32435.patch	1,63 ko	18 avril 2019	Christophe Siraut
0003-emails-validators-set-connection-timeout-32435.patch	843 octets	18 avril 2019	Christophe Siraut
0004-emails-validators-do-not-raise-validation-error-on-t.patch	1,03 ko	18 avril 2019	Christophe Siraut
0005-emails-validators-add-option-to-bypass-smtp-validati.patch	1,79 ko	18 avril 2019	Christophe Siraut
0003-emails-validators-set-connection-timeout-32435.patch	843 octets	18 avril 2019	Christophe Siraut
0002-emails-validators-address-preferred-mx-server-32435.patch	1,63 ko	18 avril 2019	Christophe Siraut
0004-emails-validators-do-not-raise-validation-error-on-t.patch	1,04 ko	18 avril 2019	Christophe Siraut
0005-emails-validators-add-option-to-bypass-smtp-validati.patch	1,79 ko	18 avril 2019	Christophe Siraut
0006-emails-validators-catch-all-dns-exceptions-32435.patch	1009 octets	18 avril 2019	Christophe Siraut