

## Passerelle - Bug #16690

### logger en mode DEBUG : 'utf8' codec can't decode ...

05 juin 2017 13:40 - Thomas Noël

|   |        |                      |                      |
|---|--------|----------------------|----------------------|
| <b>Statut:</b>  | Fermé  | <b>Début:</b>        | 05 juin 2017         |
| <b>Priorité:</b>  | Normal | <b>Echéance:</b>     |                      |
| <b>Assigné à:</b>   |        | <b>% réalisé:</b>    | 0%                   |
| <b>Catégorie:</b>   |        | <b>Temps estimé:</b> | 0:00 heure           |
| <b>Version cible:</b>   |        | <b>Planning:</b>     |                      |
| <b>Patch proposé:</b>   | Oui    |                      |                      |
| <b>Description</b>  |        |                      |                      |
| UnicodeDecodeError at /truc/machin/<br>'utf8' codec can't decode byte 0xc3 in position 4095: unexpected end of data   |        |                      |                      |
| Ça vient de passerelle/utils/__init__.py :  |        |                      |                      |
| <pre># log only if content type is allowed if content_type_match(response.headers.get('Content-Type')):     content = response.content[:settings.LOGGED_RESPONSES_MAX_SIZE]     self.logger.debug('Response Content: %r' % content,                       extra={'requests_response_content': content})</pre> |        |                      |                      |
| En effet, si le 4095ème octet est au milieu d'un caractère unicode genre «vers\xc3\xa9», alors «vers\xc3» provoque un crash lors de la conversion.  |        |                      |                      |
| Une piste pourrait être : « ... extra={'requests_response_content': '%r' % content} » ?   |        |                      |                      |
| <b>Demandes liées:</b>  |        |                      |                      |
| Lié à Passerelle - Bug #15942: erreur lors du logging des chaînes en utf-8  |        | <b>Fermé</b>         | <b>21 avril 2017</b> |

#### Révisions associées

##### Révision 6329a2a8 - 23 juin 2017 14:35 - Frédéric Péters

utils: don't break on response content cut in the middle of a character (#16690)

##### Révision 507f0994 - 23 juin 2017 14:49 - Frédéric Péters

misc: adapt test to change in requests\_response\_content logging (#16690)

#### Historique

##### #1 - 05 juin 2017 15:31 - Benjamin Dauvergne

Il faut simplement passer la chaîne de formatage principale en unicode, ensuite tout se passe généralement bien (logging se débrouille mieux avec des chaînes unicode qu'un mix ascii/utf-8).

##### #2 - 05 juin 2017 19:16 - Thomas Noël

Benjamin Dauvergne a écrit :

Il faut simplement passer la chaîne de formatage principale en unicode, ensuite tout se passe généralement bien (logging se débrouille mieux avec des chaînes unicode qu'un mix ascii/utf-8).

J'ai pas bien compris ce que ça implique ; ici c'est extra={'requests\_response\_content': content} qui pose un soucis, car plus loin dans le logeur spécifique passerelle on enregistre l'extra dans un champ JsonField, et c'est là que ça casse.

##### #3 - 06 juin 2017 00:37 - Benjamin Dauvergne

Ha ok oubliez moi. À noter que l'encodeur JSON standard accepte sans souci une str encodée en UTF-8, donc il y a un souci différent d'un simple problème d'encodage.

##### #4 - 06 juin 2017 02:00 - Thomas Noël

En fait c'est le «response.content[:settings.LOGGED\_RESPONSES\_MAX\_SIZE]» qui pose problème quand il coupe le contenu en plein milieu d'un

caractère encodé sur 16bits.

#### #5 - 23 juin 2017 14:26 - Frédéric Péters

- Fichier `0001-utils-don-t-break-on-response-content-cut-in-the-mid.patch` ajouté
- Statut changé de Nouveau à En cours
- Patch proposed changé de Non à Oui

```
- extra={'requests_response_content': content})  
+ extra={'requests_response_content': repr(content)})
```

M'irait, plutôt que '%r' % content.

Et j'ai les yeux qui saignent de cette méthode (espaces qui manquent et font des trucs très laids).

#### #6 - 23 juin 2017 14:33 - Thomas Noël

Ack

#### #7 - 23 juin 2017 14:35 - Frédéric Péters

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

```
commit 6329a2a80907d46c1cc1c3851e8450fab7cda419  
Author: Frédéric Péters <fpeters@entrouvert.com>  
Date: Fri Jun 23 14:24:05 2017 +0200
```

```
utils: don't break on response content cut in the middle of a character (#16690)
```

#### #8 - 30 juin 2017 14:28 - Josué Kouka

- Lié à Bug #15942: erreur lors du logging des chaînes en utf-8 ajouté

#### #9 - 04 août 2018 12:30 - Benjamin Dauvergne

- Statut changé de Résolu (à déployer) à Fermé

### Fichiers

---

|   |             |              |                 |
|---|-------------|--------------|-----------------|
| 0001-utils-don-t-break-on-response-content-cut-in-the-mid.patch | 1017 octets | 23 juin 2017 | Frédéric Péters |
|---|-------------|--------------|-----------------|