w.c.s. - Bug #44787

crash en fin de production d'un pdf

03 juillet 2020 11:13 - Thomas Noël

Statut:	Fermé	**Début:**	03 juillet 2020
Priorité:	Normal	**Echéance:**	
Assigné à:		**% réalisé:**	0%
Catégorie:		**Temps estimé:**	0:00 heure
Version cible:			
Patch proposed:	Oui	**Planning:**	Non

Description

Comme si libreoffice n'avait pas fait son travail :

```
Exception:
  type = '<class 'FileNotFoundError'>', value = '[Errno 2] No such file or directory: '/tmp/tmpi2e
kentd/tmphfch6u49.pdf''

Stack trace (most recent call first):
  File "/usr/lib/python3/dist-packages/wcs/wf/export_to_model.py", line 87, in transform_to_pdf
    85             raise Exception('libreoffice is failing')
    86         finally:
>   87             shutil.rmtree(temp_dir)
    88 except subprocess.CalledProcessError:
    89     transform_to_pdf = None

  locals:
     chunk = b''
     infile = <tempfile._TemporaryFileWrapper object at 0x7fcfd13cf518>
     instream = <_io.BytesIO object at 0x7fcfcb78a200>
     temp_dir = '/tmp/tmpi2ekentd'
```

Révisions associées

Révision b92ee201 - 03 juillet 2020 17:19 - Frédéric Péters

workflows: retry PDF generation if it fails for some unknown reason (#44787)

Révision 1d7e68bd - 03 juillet 2020 17:21 - Frédéric Péters

workflows: retry PDF generation if it fails for some unknown reason (#44787)

Révision e4b17b50 - 03 juillet 2020 17:46 - Frédéric Péters

misc: don't use capture_output as it requires python 3.7 (#44787)

Révision 49ab4b20 - 03 juillet 2020 17:47 - Frédéric Péters

misc: don't use capture_output as it requires python 3.7 (#44787)

Révision 8dfa5c8c - 07 juillet 2020 15:58 - Thomas Noël

workflows: isolate libreoffice UserInstallation directory (#44787)

Historique

#1 - 03 juillet 2020 14:41 - Frédéric Péters

- Fichier 0001-workflows-retry-PDF-generation-if-it-fails-for-some-.patch ajouté

- Statut changé de Nouveau à Solution proposée

- Patch proposed changé de Non à Oui

Je pose ça ici. (c'est avec du debug en plus posé sur la prod, pour voir)

#2 - 03 juillet 2020 14:59 - Frédéric Péters

Et pour info ça fonctionne :/ (cas d'appel sur la prod qui est passé lors de la deuxième itération).

#3 - 03 juillet 2020 15:15 - Frédéric Péters

- Fichier 0001-workflows-retry-PDF-generation-if-it-fails-for-some-.patch ajouté

Version qui utilise capture_output=True et "loggue" stderr.

#4 - 03 juillet 2020 17:14 - Thomas Noël

- Statut changé de Solution proposée à Solution validée

Tristesse, mais on n'arrive pas à en savoir plus, alors go.

#5 - 03 juillet 2020 17:20 - Frédéric Péters

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

```
commit b92ee201e21506fd07032010bc638c4a911ab19c
Author: Frédéric Péters <fpeters@entrouvert.com>
Date:   Fri Jul 3 14:37:38 2020 +0200

    workflows: retry PDF generation if it fails for some unknown reason (#44787)
```

#6 - 03 juillet 2020 17:44 - Frédéric Péters

Trop à la page, capture_output c'est uniquement à partir de 3.7...

#7 - 03 juillet 2020 20:16 - Frédéric Péters

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

#9 - 07 juillet 2020 14:12 - Thomas Noël

Exemple de crash loggué sur wcs.node1 (Mon, 06 Jul 2020 16:34:59 +0200) :

```
/var/lib/wcs/demarches.strasbourg.eu /tmp/tmp8hlg03w1 /tmp/tmp8hlg03w1/tmpfdakn8yt
 -> Error0: /var/lib/wcs/demarches.strasbourg.eu stdout:b'' stderr:b''
 -> Error1: /var/lib/wcs/demarches.strasbourg.eu stdout:b'' stderr:b''
 -> Error2: /var/lib/wcs/demarches.strasbourg.eu stdout:b'' stderr:b''
 -> Output2: /var/lib/wcs/demarches.strasbourg.eu stdout:b'' stderr:b''
 -> Finally: /var/lib/wcs/demarches.strasbourg.eu /tmp/tmp8hlg03w1 []
```

stdout et stderr vides :/

Pas de dmesg ou journalctl qui remonte hier à 16h34

#12 - 07 juillet 2020 15:30 - Thomas Noël

Je pense que j'ai trouvé la source du pépin. En lançant 5 processus de conversion **concurrents** sur ma machine, la plupart du temps 2 ou 3 se terminent mal et sans aucune sortie. On aurait donc un problème de concurrence. En lançant l'affaire via "strace" je constate que libreoffice reécrit par exemple /home/thomas/.config/libreoffice/4/user/registrymodifications.xcu

Je n'ai pas trouvé d'option dans libreoffice pour lui force un emplacement de .config ailleurs.

En bidouillant avec un HOME=/tmp/random avant chaque lancement, ça semble mieux passer.

#13 - 07 juillet 2020 15:49 - Thomas Noël

- Fichier 0001-workflows-isolate-libreoffice-UserInstallation-direc.patch ajouté

- Statut changé de Solution déployée à Solution proposée

En creusant, je suis arrivé sur https://ask.libreoffice.org/en/question/7284/multiple-instances-for-c-automation/ et l'usage de « -env:UserInstallation=file:///tmp/xxx »

Ma proposition, donc, en gardant quand même pour l'instant la sécurité moche de "tenter 3 fois".

#14 - 07 juillet 2020 15:52 - Christophe Siraut

- Statut changé de Solution proposée à Solution validée

#16 - 07 juillet 2020 16:09 - Thomas Noël

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

```
commit 8dfa5c8c62d21d124ae1b0bdbe4da0cd31037925
Author: Thomas NOEL <tnoel@entrouvert.com>
Date:   Tue Jul 7 15:42:42 2020 +0200

    workflows: isolate libreoffice UserInstallation directory (#44787)
```

#18 - 07 juillet 2020 16:16 - Frédéric Péters

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

#19 - 07 juillet 2020 17:13 - Benjamin Dauvergne

Il me semble que les gens qui utilisent loffice headless le lancent en mode serveur pour éviter ça.

#20 - 15 septembre 2022 16:20 - Pierre Ducroquet

Je plussois le commentaire de Benjamin, je suis tombé sur plusieurs instances de LibreOffice tournant sur un wcs et j'ai été un peu surpris.
Outre la possibilité de mieux comprendre ce qu'il se passe, ça nous apporterait aussi un gain en performance.
Sur ma machine, convertir un document en PDF avec la commande de wcs prend 900ms/1s. Si par contre je lance LibreOffice en headless (il pourrait alors tourner avec un utilisateur à part, ou être relancé automatiquement par systemd ou que sais-je, avec en prime le --safe-mode) et que je m'y connecte en Python, la conversion prend 300ms.

Pour référence, le lancement de libreoffice:

```
libreoffice --headless --safe-mode --accept="socket,host=localhost,port=2002;urp;StarOffice.ServiceManager"
```

Et le code python:

```
#!/usr/bin/env python3

import uno
# no joke...
from com.sun.star.beans import PropertyValue

def from_local():
    localContext = uno.getComponentContext()
    resolver = localContext.ServiceManager.createInstanceWithContext("com.sun.star.bridge.UnoUrlResolver", loc
alContext)
    context = resolver.resolve("uno:socket,host=localhost,port=2002;urp;StarOffice.ComponentContext")
    desktop = context.ServiceManager.createInstanceWithContext("com.sun.star.frame.Desktop", context)
    component = desktop.loadComponentFromURL("file:///home/pierre/Documents/demo.odt", "_default", 0, ())
    pv = PropertyValue()
    pv.Name = "FilterName"
    pv.Value = "writer_pdf_Export"
    component.storeToURL("file:///tmp/test01.pdf", [pv])
    component.close(0)
```

My deux centimes.

#21 - 15 septembre 2022 17:10 - Benjamin Dauvergne

Ouvre un autre ticket, personne ne lit les vieux tickets.

#22 - 19 septembre 2022 19:53 - Frédéric Péters

Personne ne lit les tickets **fermés**. (ce qui n'est même pas vrai).

Fichiers

0001-workflows-retry-PDF-generation-if-it-fails-for-some-.patch	2,03 ko	03 juillet 2020	Frédéric Péters
0001-workflows-retry-PDF-generation-if-it-fails-for-some-.patch	2,09 ko	03 juillet 2020	Frédéric Péters
0001-workflows-isolate-libreoffice-UserInstallation-direc.patch	1,08 ko	07 juillet 2020	Thomas Noël