

## Thot - Tâche #28562

Scénario # 28733 (Terminé (Sprint)): Thot et import AAF au format ENT2VA: nombreuses erreurs

### Thot: echec import AAF : doublons dans les classes d'un établissement donné (dans Idif)

06/07/2019 03:18 PM - arnaud grossir

<b>Status:</b>	Fermé	<b>Start date:</b>	06/07/2019
<b>Priority:</b>	Haut	<b>Due date:</b>	
<b>Assigned To:</b>	Joël Cuissinat	<b>% Done:</b>	100%
<b>Target version:</b>	sprint 2019 27-35 Equipe MENSUR (été)	<b>Estimated time:</b>	0.00 hour
		<b>Spent time:</b>	0.00 hour
<b>Description</b>			
<p>Bonjour,</p> <p>Nous avons un serveur Thot version 2.6.2 L'import de l'AAF se fait actuellement en mode 'aaf-complet-maj'</p> <p>Nous remarquons depuis quelques jours que cet import ne se fait pas correctement, par exemple nous avons des erreurs de type:</p> <pre>modifying entry "ENTStructureUAI=0251556Fldap_modify: Type or value exists (20)   additional info: ENTStructureClasses: value #0 provided more than once ,ou=structures,dc=enoe,dc=net"</pre> <p>Après recherche, il se trouve que certains établissements présent dans les xml issus de l'AAF présentent des doublons de classe</p> <p>Ceci est dû au fait que ces établissements ont fait leur bascule STS pour l'année N+1 en demandant la récupération des classes qui existent déjà</p> <p>Or, le traitement AAF ne fait pas le distingo entre l'année en cours et l'année en préparation, ce qui fait qu'il récupère toutes les classes des deux périodes. Il génère donc des établissements avec une entrée 'ENTStructureClasses' qui possède des valeurs en double</p> <p>exemple (retours à la ligne pour + de visibilité):</p> <pre>&lt;attr name="ENTStructureClasses"&gt; &lt;value&gt;31\$troisieme 1\$10310019110&lt;/value&gt; &lt;value&gt;32\$troisieme 2\$10310019110&lt;/value&gt; &lt;value&gt;33\$troisieme 3\$10310019110&lt;/value&gt; &lt;value&gt;34\$troisieme 4\$10310019110&lt;/value&gt; &lt;value&gt;41\$quatrieme 1\$10210001110&lt;/value&gt; &lt;value&gt;42\$quatrieme 2\$10210001110&lt;/value&gt; &lt;value&gt;43\$quatrieme 3\$10210001110&lt;/value&gt; &lt;value&gt;44\$quatrieme 4\$10210001110&lt;/value&gt; &lt;value&gt;45\$quatrieme 5\$10210001110&lt;/value&gt; &lt;value&gt;51\$cinquieme 1\$10110001110&lt;/value&gt; &lt;value&gt;52\$cinquieme 2\$10110001110&lt;/value&gt; &lt;value&gt;53\$cinquieme 3\$10110001110&lt;/value&gt; &lt;value&gt;54\$cinquieme 4\$10110001110&lt;/value&gt; &lt;value&gt;55\$cinquieme 5\$10110001110&lt;/value&gt; &lt;value&gt;61\$sixieme 1\$10010012110&lt;/value&gt; &lt;value&gt;62\$sixieme 2\$10010012110&lt;/value&gt; &lt;value&gt;63\$sixieme 3\$10010012110&lt;/value&gt; &lt;value&gt;64\$sixieme 4\$10010012110&lt;/value&gt; &lt;value&gt;65\$sixieme 5\$10010012110&lt;/value&gt;  &lt;value&gt;31\$troisieme 1\$10310019110&lt;/value&gt; &lt;value&gt;32\$troisieme 2\$10310019110&lt;/value&gt; &lt;value&gt;33\$troisieme 3\$10310019110&lt;/value&gt;</pre>			

```
<value>34$troisieme 4$10310019110</value>
<value>35$troisieme 5$10310019110</value>
<value>41$quatrieme 1$10210001110</value>
<value>42$quatrieme 2$10210001110</value>
<value>43$quatrieme 3$10210001110</value>
<value>44$quatrieme 4$10210001110</value>
<value>45$quatrieme 5$10210001110</value>
<value>51$cinquieme 1$10110001110</value>
<value>52$cinquieme 2$10110001110</value>
<value>53$cinquieme 3$10110001110</value>
<value>54$cinquieme 4$10110001110</value>
<value>55$cinquieme 5$10110001110</value>
<value>61$sixieme 1$10010012110</value>
<value>62$sixieme 2$10010012110</value>
<value>63$sixieme 3$10010012110</value>
<value>64$sixieme 4$10010012110</value>
<value>65$sixieme 5$10010012110</value>
</attr>
```

C'est pour cela que l'import des fichiers issus de l'AAF dans l'annuaire du ldap Thot échoue (vous remarquerez que le premier bloc de classes est celui de l'année en cours et le second celui de l'année à venir, avec une classe supplémentaire: la 3ème 5)

J'ai vérifié côté Scribe, cela se passe très bien, un rapide coup d'oeil aux scripts python me fait penser que l'import AAF des scribe détecte s'il y a des classes en doublon ou pas (il reçoit les mêmes fichiers, les doublons y sont bien présent, mis à part qu'il ne reçoit que les données qui le concerne, et non un ensemble de données provenant de plusieurs établissement comme pour Thot).

Pour reproduire le problème, il suffit de mettre une classe en double pour un établissement dans le fichier AAF des établissements: **Complet\_EtabEducNat\_0000.xml** (a l'intérieur de l'attribut 'ENTStructureClasses')

Pour pallier rapidement au problème j'ai:

1- désactivé la re-écriture des template dans le script /usr/sbin/aaf-complet-maj (nous ne fonctionnons qu'avec ce mode, nous n'utilisons pas aaf-delta)

j'ai donc mis en remarque la ligne "#usr/sbin/gentemplate.py delta"

2- re-écrit le template Cheetah /var/lib/eole/aaf/maj\_etablissement.ldif (avec mes connaissances quasi nulles sur le sujet, mais bon...).

Je sais que pour faire plus propre je n'aurais pas du désactiver la ligne dans aaf-complet-maj et aurais dû plutôt modifier le script gentemplate.py, mais c'est + compliqué à réaliser... j'ai été au + simple avant de vous faire part de ce problème

Donc, dans maj\_etablissement.ldif, j'ai remplacé:

```
%if %%varExists('ENTStructureClasses')
replace: ENTStructureClasses
%for %%i in %%ENTStructureClasses.split("\n")
%if %%i != ''
ENTStructureClasses: %%i
%end if
%end for
-
%end if
```

par:

```
%if %%varExists('ENTStructureClasses')
%set test_Classes = {}
replace: ENTStructureClasses
%for %%i in %%ENTStructureClasses.split("\n")
```

```
%if %%i != ''
%if test_Classes.has_key(%%i)
%continue
%else
%set test_Classes[%%i] = '1'
ENTStructureClasses: %%i
%end if
%end if
%end for
-
%end if
```

Ca fonctionne, le fichier delta\_update.ldif ne comporte plus de doublons, mais ce n'est pas propre et il y a sans doute des subtilités qui m'échappent.

(de plus ça a le défaut de ne pas supprimer les NOUVELLES classes de l'année en préparation... mais à la limite c'est pas très grave et je crois qu'on ne saura jamais gérer ce problème)

Donc, je me demande s'il serait possible de votre côté de regarder pour empêcher l'apparition de ces doublons dans les ldif générés par aaf-complet, aaf-complet-maj, et/ou aaf-delta

Je me demande aussi si le mieux ne serait pas de faire ce travail en amont, lors du nourrissage de la base mysql 'eoleaaf', et de vérifier les doublons pour les classes, mais aussi pour les groupes...

Toutes les académies sont certainement concernées, et chaque établissement, durant la période de préparation de ses structures (divisions et groupes) est susceptible de générer ce problème.

J'ai pu avoir, par l'intermédiaire d'un collègue, cet échange (signalements STS vis à vis de l'aaf): cela concerne les services, mais, de manière indirecte, les divisions et les groupes

Bonjour à tous(tes),  
Dans le fichier "EtabEducNat" la balise "ENTStructureClasses" comporte les classes de l'année en cours (2016-2017) et celles de l'année à venir (2017-2018) puisque les services enseignants 2016-2017 restent effectifs jusqu'au 31 août alors que la partie scolarité des élèves est déjà celle de l'année à venir.  
Un fournisseur ENT me demande si, pour cette période, il peut afficher uniquement les services de l'année en cours.

Réponse

Le fournisseur ENT peut faire ce qu'il veut. Par contre, je ne vois pas comment il va différencier les services en cours et les services à venir.... dans les exports, il n'y a pas de distinction. La seule chose peut être c'est qu'ils ont une date de création de leur enregistrement donc celui qui a la date ancienne, c'est affectation en cours et celui qui a une date récente, c'est affectation à venir.  
Cet accord relève du fournisseur (capacité à le faire ou pas) et du choix de votre académie. Pas de directive ou contrainte de ce côté là.

Donc j'ai l'impression qu'on ne doit pas s'attendre à des modifications des script d'extraction AAF, malheureusement...

Merci par avance de l'attention que vous porterez à ma demande

Cordialement

## History

#1 - 06/14/2019 11:35 AM - Luc Bourdot

- Tracker changed from Demande to Anomalie

- Priority changed from Normal to Haut

Merci pour ce retour très détaillé

**#2 - 07/11/2019 10:51 AM - Joël Cuissinat**

- Parent task set to #28733

**#3 - 07/16/2019 04:23 PM - Joël Cuissinat**

- Subject changed from *Thot: echec import Aaf : doublons dans les classes (dans Idif)* to *Thot: echec import Aaf : doublons dans les classes d'un établissement donné (dans Idif)*

**#4 - 07/16/2019 04:31 PM - Joël Cuissinat**

Dans les fichiers que nous possédons, c'est le cas pour "ENTStructureUAI=0940072T", les 12 classes sont clairement en double et cela vient directement du fichier XML, on pourrait donc gérer ça dès la lecture...

Les attributs ENTStructureClasses et ENTStructureGroupes sont ceux qui peuvent être concernés.

Après plusieurs tests, je propose de supprimer les doublons directement dans `util.parse_multi()`.

Cela devrait régler le problèmes des prénoms en double en même temps ;)

**#5 - 07/16/2019 04:50 PM - Joël Cuissinat**

- Status changed from *Nouveau* to *En cours*

**#6 - 07/18/2019 11:23 AM - Joël Cuissinat**

- Status changed from *En cours* to *Résolu*

- Assigned To set to *Joël Cuissinat*

- % Done changed from 0 to 100

**#7 - 09/12/2019 03:52 PM - arnaud grossir**

Bonjour, ça marche bien.

Merci  
Cordialement

**#8 - 09/17/2019 09:15 AM - Joël Cuissinat**

- Status changed from *Résolu* to *Fermé*

- Remaining (hours) set to 0.0