

Micro projet - Analyse des données hétérogènes

Travail effectué par : CHANI Yassine, JGUIRIM Baha Eddine, ZHARIGELE

I. Contexte

Hop3x est un environnement de travaux pratiques pour l'apprentissage de la programmation. Il permet au tuteur de suivre, à distance, l'activité de programmation des apprenants en lui fournissant, d'une part, des informations qualitative (indicateurs) sur cette activité, et d'autre part, une visualisation en temps réel des productions (code source des programmes).

Notre projet consiste donc à utiliser ces traces pour calculer quelques indicateurs qui correspondent à chaque apprenant.

II. Travail réalisé

1. Talend

Le travail réalisé se divise en deux parties, la première partie effectuée sur le logiciel Talend dans laquelle nous avons produit un « job » qui prend en entrée les fichiers XML et les transforme en données sql dans une base MySQL.

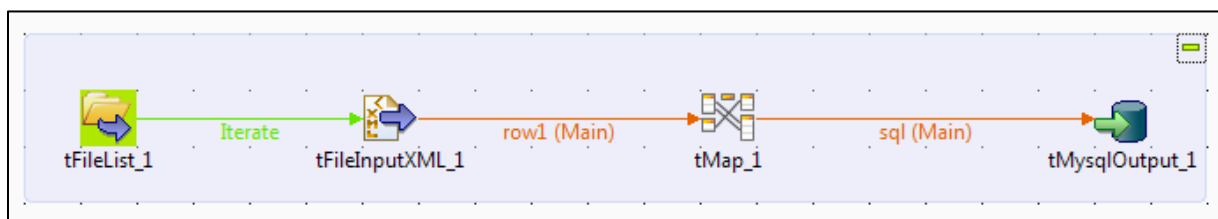


Figure 1 : vue globale de job réalisé en Talend

Cette transformation n'est pas une simple conversion de XML vers MySQL, en effet, les traces enregistrées dans les fichiers XML sont très volumineuses, car on stocke toutes les actions produites par l'étudiant durant le TP. Donc nous avons sélectionné seulement les balises qu'on aura besoin, à savoir :

- **nom** et **id** de la balise **auteur** : nom et l'identifiant de l'étudiant en question

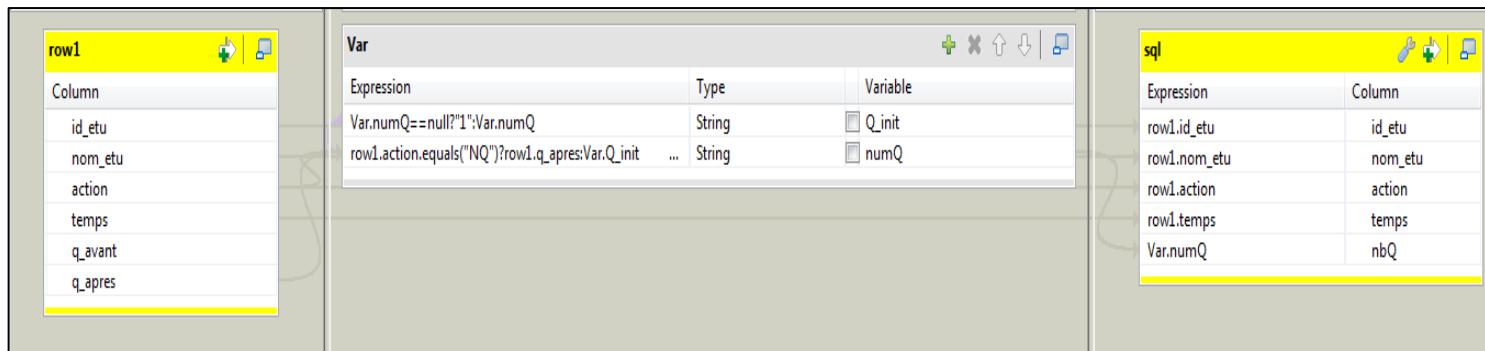
⇒ Remarque : cette balise n'existe pas initialement dans le fichier xml, nous l'avons ajoutée pour bien identifier chaque étudiant dans la base de données)

- **K** : action effectuée par l'étudiant, nous avons pris seulement les actions suivantes **NQ** : question suivante, **CAE** : compilation automatique avec erreurs, **E** : compilation sans erreurs, **CME** : compilation manuelle avec erreurs)

⇒ NQ, QAPRES et QAVANT ont été utilisées pour obtenir, pour chaque action, le numéro de la question (voir figure 3)

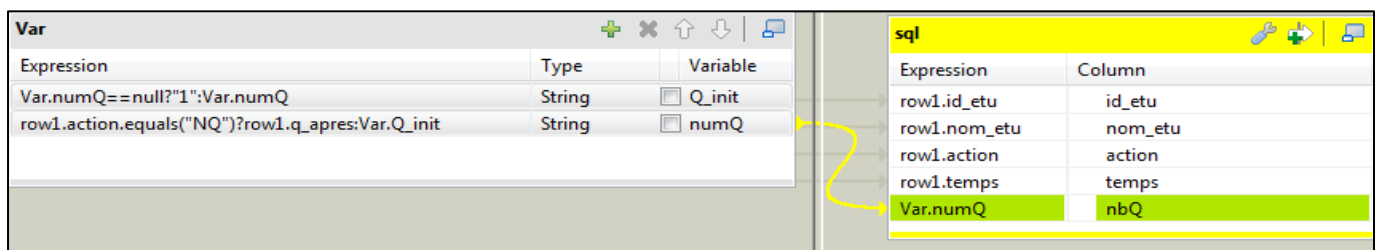
- **T** : le temps de chaque action

Pour transformer les données xml en données sql on doit faire une « Map » qui traduit les relations entre ces données.



row1	Var	sql
id_etu	Expression: Var.numQ==null?1:Var.numQ, Type: String, Variable: Q_init	row1.id_etu
nom_etu	Expression: row1.action.equals("NQ")?row1.q_apres:Var.Q_init, Type: String, Variable: numQ	row1.nom_etu
action		row1.action
temps		row1.temps
q_avant		Var.numQ
q_apres		nbQ

Figure 2 : vue globale de la map dans le job de Talend



Var	sql
Expression: Var.numQ==null?1:Var.numQ, Type: String, Variable: Q_init	row1.id_etu
Expression: row1.action.equals("NQ")?row1.q_apres:Var.Q_init, Type: String, Variable: numQ	row1.nom_etu
	row1.action
	row1.temps
	Var.numQ
	nbQ

Figure 3 : calculer le numéro de la question à l'aide de 2 variables et les deux balises NQ et QAPRES

Explication de la figure 3 : La colonne nbQ dans la base de donnée prend comme valeur le numéro de chaque question, cette valeur a été calculée depuis <NQ> et <QAPRES> comme le suivant :

Initialement le numéro de la question est 1 (variable Q_init = 1 initialement , puis le numéro de la question courante)

Ensuite si l'action est NQ (donc l'étudiant a changé la question) on récupère la valeur de la balise QAPRES qui contient le numéro de la question suivante, sinon on garde le même numéro de la question

2. Interface Web

Pour calculer et afficher les indicateurs nous avons développé une interface en PHP. Cette interface contient une première page qui liste tous les étudiants, avec une possibilité de chercher un étudiant à partir de son identifiant ou son nom (voir figure 4).

Micro-projet Analyse des données hétérogènes

Hop3X : affichage des indicateurs

10 records per page

Search:

Indicateurs par étudiant

Indicateurs globales en graphe

id etudiant	Nom etudiant	numéro TP	Date	Afficher détails
123401	PIERRE FOUCAULT	tp1	29-03-2013	<a>Afficher
123402	Pierre JACOBONI	tp1	27-03-2013	<a>Afficher
123403	PIERRE JOUANNET-VAN DER SANDE	tp1	03-04-2013	<a>Afficher
123404	PIERRE-FRANCOIS NAVEAU	tp1	29-03-2013	<a>Afficher
123405	QUENTIN BRUEL	tp1	27-03-2013	<a>Afficher

Figure 4 : page d'accueil

Depuis cette page on a la possibilité d'afficher des indicateurs graphiques globales sur tout le groupe (voir figure 5), ou en affichant les indicateurs propres à chaque étudiant (voir figure 6).

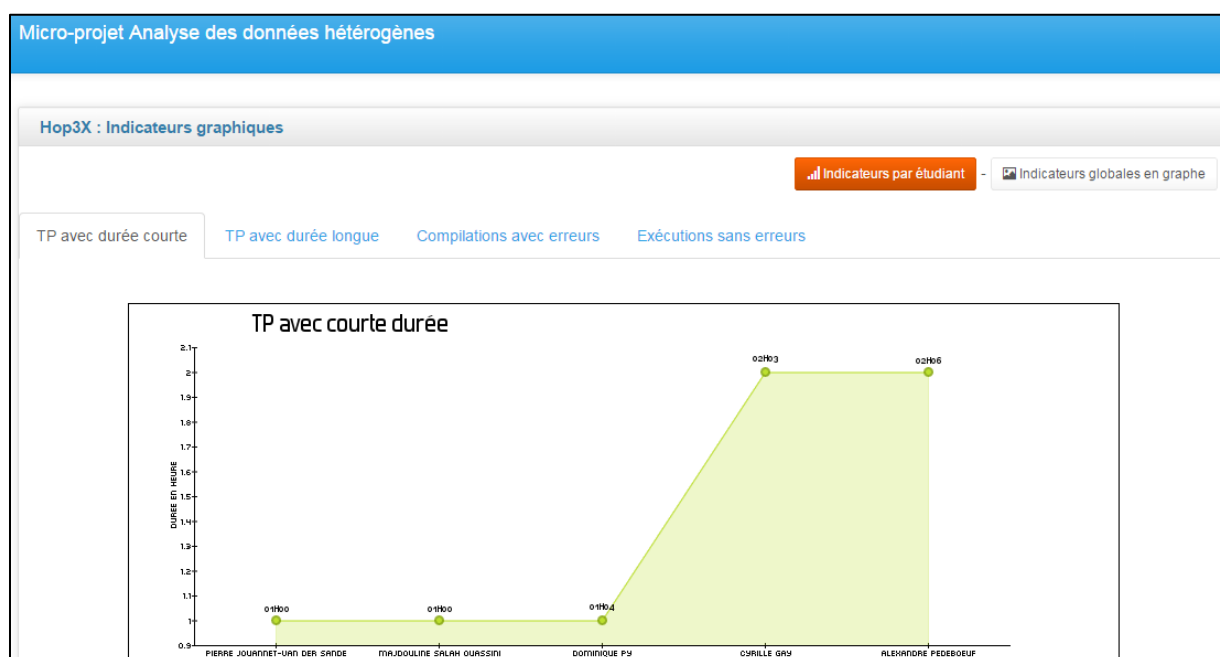


Figure 5 : page affichant les indicateurs graphique pour tout le groupe

Cette page affiche des indicateurs graphiques en fonction de tout le groupe, on distingue comme autres indicateurs présentés:

- Les étudiants ayant passés plus de temps dans leurs TP
- Les étudiants ayant passés moins de temps dans leurs TP
- Les étudiants ayant compilés avec plus d'erreurs
- Les étudiants ayant compilés avec succès

On peut aussi consulter les indicateurs liés à chaque étudiant en cliquant sur le bouton « afficher » dans la ligne qui contient le nom de l'étudiant, on obtient alors une liste des indicateurs propre à l'étudiant sélectionné (voir figure 6).

Micro-projet Analyse des données hétérogènes

Informations sur l'étudiant : ANTHONY BLANQUART

[← Retour à l'accueil](#)

N° TP	N° Question	Temps de début	Executions avec erreurs ⚠	Executions sans erreurs ✔	Dernière compilation
tp1	1	28-03-2013 09:36	15	1	succès
tp1	2	28-03-2013 10:01	14	1	succès
tp1	3	28-03-2013 10:23	1	3	succès
tp1	4	28-03-2013 10:38	5	1	échec
Total	4	Durée de TP 02H47	35	6	4 (3 ✔ + 1 ⚠)

Figure 6 : page affichant les indicateurs propres à chaque étudiant

Cette page contient des indicateurs pour chaque étudiant qui sont :

- La durée de TP
- Le temps de début de chaque question
- Le nombre d'exécution avec erreurs
- Le nombre d'exécutions avec succès
- Le nombre des questions abordées au cours de TP
- Le nombre des questions finissant par une exécution avec succès
- Le nombre des questions finissant par une exécution avec erreurs