

Débuts prometteurs pour l'ETL de Talend

INTÉGRATION. La version 1.0 de la plate-forme open source de l'éditeur français est téléchargeable depuis quelques semaines. Objectif : asticoter Informatica et IBM.

Avec 10 000 téléchargements et plus d'une dizaine de promesses de souscription en un mois, le lancement début octobre de l'ETL open source de Talend a tenu toutes ses promesses. La disponibilité, d'ici à une semaine, de la version 1.1 de la solution de l'éditeur parisien devrait encore favoriser son essor. Cette nouvelle mouture enrichit son designer de flux et son moteur de transformation (basés sur RCP) d'un outil d'administration des scénarios d'extraction. Ces trois composants, Talend les offre à la communauté open source. La jeune pousse se finance sur le conseil, la formation, et la maintenance autour de sa plate-forme.

A l'origine, Talend vise les projets d'intégration de milieu de gamme. Avec une mention spéciale pour ceux n'intégrant aucune dimension décisionnelle (migration et synchronisation de données, par exemple). Pour autant, face aux ETL d'entreprise d'Informatica et d'Ascential/IBM, le Français vante ses atouts technologiques – notamment ceux liés

aux connecteurs. Sa plate-forme en propose nativement une quarantaine, surtout vers des bases de données du marché. Mais elle accepte aussi les milliers d'autres connecteurs développés par la communauté Perl (un langage optimisé pour l'extraction d'informations de fichiers), destinés au monde open source.

Des connecteurs dénoncés par les rivaux. Chez les concurrents, c'est précisément les connecteurs de Talend que l'on montre du doigt. « Ils ne sont certifiés par aucun éditeur – SAP, Hyperion ou Salesforce, par exemple. Or, dans les projets d'intégration de données, les entreprises soumettent leur fournisseur à une obligation de résultat. Autrement dit, à un engagement sur la stabilité de la plate-forme », explique Didier Guyomarc'h, d'Informatica. On voit cependant mal la start up obtenir, en un an d'existence, une certification SAP : le processus dure au moins trois ans, et s'avère très coûteux. « Les entreprises développant leur propre connecteur pour synchroniser SAP et Siebel sont légion. Ont-elles besoin d'une certification SAP ?

Encore peu de projets ETL en open source

- **Pentaho** : éditeur américain couvrant toute la chaîne décisionnelle : reporting, analyse, datamining et workflow. En avril dernier, il rachète le projet Kettle, un ETL open source assurant des fonctions de fédération virtuelle de données (EII). Pentaho dispose d'environ 10 millions d'euros d'investissement.
- **Clover.ETL** : sponsorisé par Berlios et Opensys. ETL développé en Java et dédié à la transformation des

données structurées. Fourni avec un designer conçu sous Eclipse. Peut être utilisé indépendamment ou embarqué dans les applications.

- **Enhydra Octopus** : projet soutenu par le consortium Objectweb. Octopus repose sur le serveur d'applications Enhydra. Conçu en Java, l'ETL se connecte aux sources JDBC. Les transformations sont définies dans un fichier XML [génération de code Javascript].

CE QU'IL EN PENSE



Geercken Uwe,
responsable Business
Intelligence chez
Swissport

« Trois questions restent en suspens.

Nous avons téléchargé l'ETL de Talend dans le cadre d'un projet décisionnel en cours de réalisation. Pour l'instant, celui-ci implique un petit datawarehouse – bâti sur une base MySQL – que nous alimentons manuellement, interrogé par l'outil de reporting d'Hyperion. Nous avons vite constaté que l'ETL du Français était très bien architecturé pour la transformation ou la fusion de données. Mais trois questions restent en suspens. L'une a trait à l'ordonnancement des tâches, aujourd'hui possible pour les seuls petits volumes. L'autre concerne l'administration des scénarios d'alimentation, encore très réduite. La dernière porte sur la qualité de données. Une fonction que Talend devrait rendre disponible en début 2007. Nous déciderons, d'ici deux à trois mois, si nous allons ou non souscrire à leur offre de support. »

Pas le moins du monde. Nous visons précisément ce type d'organisations », rétorque Bertrand Diart, PDG de Talend.

Testé par Caggemini. Au moins trois briques fonctionnelles manquent encore à Talend pour rivaliser avec Informatica ou IBM : la qualité de données, l'ordonnancement de tâches et la gestion centralisée des métadonnées. Mais ce décalage n'a pas empêché le jeune éditeur de se rapprocher des grands noms du service, tels Sopra, Bull ou Caggemini. Ce dernier teste, d'ailleurs, en ce moment, le potentiel de Talend chez un grand client.

Avec 3 millions d'euros de liquidités, la start up peut tenir vingt-quatre mois dans l'hypothèse d'une absence de revenus. L'éditeur mise sur un chiffre d'affaires avoisinant le million d'euros en 2007. ●

VINCENT BERDOT

v.berdot@01informatique.presse.fr