

## DesEsper: un logiciel de pré-traitement de flux appliqué à la surveillance des centrales hydrauliques

Frédéric Flouvat \*, Sébastien Gassmann \*\*, Jean-Marc Petit \*\*, Alain Ribière \*\*\*\*

\* Université de la Nouvelle-Calédonie, PPME, F-98851, Nouméa, [frederic.flouvat@univ-nc.nc](mailto:frederic.flouvat@univ-nc.nc)

\*\* INSA Lyon, LIRIS, UMR5205 CNRS, F-69621, Villeurbanne, [prenom.nom@insa-lyon.fr](mailto:prenom.nom@insa-lyon.fr)

\*\*\* EDF R&D, département STEP, F-78401, Chatou, [alain.ribiere@edf.fr](mailto:alain.ribiere@edf.fr)

Ces dernières années, la gestion des flux de données est devenue une thématique importante en base de données. En effet, un nombre croissant d'applications sont fondées sur la génération et le traitement de flux (ou flots) d'événements, i.e. de séries temporelles ordonnées d'événements hétérogènes et distribués arrivant quasiment en temps réel. Le processus de stocker les données pour ensuite les traiter ne suffit plus. Un traitement "à la volée" des informations doit être possible. Dans le cadre de son activité, EDF génère également de grandes quantités d'informations sous la forme de flux de données, notamment pour la surveillance des centrales hydrauliques. Actuellement, leur exploitation est fastidieuse car "manuelle" dans la quasi-totalité des sites EDF. De plus, les capacités de prévention des pannes sont limitées en raison du manque d'informations agrégées sur le fonctionnement et l'évolution des systèmes observés.

Dans ce contexte, nous avons développé un logiciel de pré-traitement de flux de données : *DesEsper*. Son objectif est de filtrer, de restructurer et de rediriger des flux d'événements hétérogènes et distribués vers des logiciels d'analyse tierces ou plus simplement vers des outils de visualisation (p.ex. tableur Excel). Le système à base règles de *DesEsper* permet de filtrer finement les événements par l'intermédiaire de contraintes classiques de type "*select from where*", de contraintes temporelles sur les événements et d'expressions régulières sur leur description. Contrairement aux autres solutions, *DesEsper* traite les événements en fonction de leur date réelle d'occurrence, i.e. celle inscrite par la source. A notre connaissance, les systèmes existants traitent les événements en fonction de leur date d'arrivée dans le système de gestion d'événements et non pas en fonction de leur date réelle, ce qui peut totalement fausser l'analyse. *DesEsper* intègre ainsi la nature temporelle et textuelle des événements au coeur de son architecture. Il permet également de formater totalement, et de manière déclarative, le flux en sortie en utilisant pour cela des notations issues des bases de données classiques (p.ex. pour le format des dates). Pour finir, *DesEsper* traite les éventuels retards au niveau des événements, en réinjectant dans le moteur de règles les événements censés arriver après, ce qui permet de détecter potentiellement de nouvelles séquences d'événements.

De part son architecture, *DesEsper* est un système de gestion de flux particulièrement adapté à des applications où les règles sont liées à des informations critiques (p.ex. des alarmes dans les centrales hydrauliques), et où l'environnement est relativement fiable (peu d'événements en retard), ce qui était le cas chez EDF. Il a notamment été utilisé avec succès pour étudier l'évolution dans le temps de la séquence de démarrage et d'arrêt de groupes d'une centrale hydraulique.