



SHAMAN : le projet européen pour la préservation de la mémoire informatique

par Manon Vincent , le 25/06/2012 à 14:38

Projet collaboratif issu de l'association de plusieurs sociétés et universités européennes, le projet SHAMAN qui visait à trouver une solution pour la préservation des données numériques sur le long terme vient de publier son rapport final, ayant atteint son objectif fin de l'année 2011.

Lancé en 2007 et partiellement financé par l'Union européenne (à hauteur de 8,4 millions sur 12,3 millions au total), le projet **SHAMAN** (pour Sustaining Heritage Access through Multivalent ArchiviNg), vient d'éditer son rapport final après avoir atteint son objectif fin 2011. Comme Ruben Riestra, coordinateur du projet, le rappelle, le but de SHAMAN qui était de résoudre le problème de la perte de mémoire informatique a été atteint grâce à une « structure qui fait de la conservation numérique une réalité pour quasiment n'importe quel format de données ». Grâce à la structure mise en œuvre, les données seront donc préservées, mais aussi « lisibles, accessibles et utilisables par les générations futures ». Ce projet visait à avoir une approche globale du problème de conservation des données, afin de permettre notamment l'intégration de **l'architecture SHAMAN** (ou SRA, pour **SHAMAN** Reference Architecture) dans celle de n'importe quelle organisation. Le **framework SHAMAN** propose des outils d'analyse, de gestion, de contrôle d'accès ainsi que des outils pour la réutilisation des données via différentes archives et bibliothèques documentaires. Au niveau de l'hébergement des données, la technique retenue est le **Grid Computing** (1), car comme Ruben Riestra l'a souligné : « Pour le futur, nous avons besoin de solutions robustes, sur le long terme, qui peuvent sécuriser des données et métadonnées dans de nombreux formats ». De sorte que la technique du Cloud Computing n'a pas été retenue en raison de l'imaturité des technologies : comme le note Ruben Riestra, « dans la préservation des données électroniques, l'ennemi est le temps ».

Trois prototypes du projet ont déjà été développés :

- un a été déployé en association avec la bibliothèque nationale allemande,
- un deuxième a été déployé dans un contexte industriel ce qui a permis d'améliorer l'efficacité des processus ainsi que les temps de sauvegarde,
- un troisième prototype a été testé dans le domaine des sciences. Ce domaine est d'ailleurs particulièrement intéressant à étudier puisqu'il génère un grand nombre de données qu'il est nécessaire de pouvoir classer et consulter au fil du temps.

Il semble, en outre, qu'avec le développement de l'open data (ouverture des données), ce projet trouve facilement de nouvelles applications.

(1) Le Grid Computing aussi appelé calcul partagé ou calcul distribué, est une technique de calcul consistant à partager les ressources d'un grand nombre d'ordinateurs de type souvent très différents, disséminés dans des lieux éventuellement très éloignés, et reliés par Internet ou d'autres types de réseaux spécialisés.

Source : *01net.com*, Pierre Fontaine.

Catégories : Brève, Archivage, Archivage électronique, France, Europe (hors France)

Tags : archivage électronique, données numériques, Grid Computing, SHAMAN