

Les bases de données XML

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 16 mars 2006

<https://www.bortzmeyer.org/base-donnees-xml.html>

Les bases de données existent en informatique depuis de très nombreuses années et, aujourd'hui, presque toutes celles qui sont en production sont des bases de données relationnelles. Les bases de données XML vont-elles les remplacer, au moins partiellement ?

Une base de données XML est une base de données qui stocke des documents XML. Plus précisément (car un SGBD relationnel peut stocker des documents XML comme texte), elle est optimisée pour le XML et permet d'utiliser des langages de requête XML comme XPath ou XQuery).

Voici un exemple d'utilisation d'une base de données XML ultra-simple, nommée Agouti en l'honneur d'un mignon mammifère (le code sources est disponible (en ligne sur <https://www.bortzmeyer.org/files/agouti.py>)) :

```
% ./agouti init experimental
% ./agouti add experimental test.xml
Added test.xml as ID 1
% ./agouti query experimental '//*[text()="Moi"]'
1: <author>Moi</author>
% ./agouti delete experimental 1
% ./agouti query experimental '//*[text()="Moi"]'
%
```

Ce qui manque à agouti ?

- Indexation : pas seulement du texte ou du nom des éléments mais aussi des relations entre éléments pour des expressions comme `foo/bar`. Le problème est très complexe et a fait l'objet de nombreuses publications comme l'excellente "*Accelerating XPath Evaluation in Any RDBMS*" <<http://www.inf.uni-konstanz.de/dbis/publications/download/accelerating-locsteps.pdf>>.
- Validation : tester les documents en entrée contre un schéma.
- Méta-information : nom du document, date de dépôt.

Alors, quels sont les logiciels sérieux de base de données XML ? Je n'ai testé que DB-XML (les autres sont en Java et ne semblent pas tourner sur du logiciel libre) et cela ne me semblait pas encore très au point.

- DB-XML <<http://dev.sleepycat.com/documentation/bdbxml.html>> (Sleepycat <<http://www.sleepycat.com/>>). Voir aussi un bon article de XML.com <<http://www.xml.com/pub/a/2003/05/07/bdb.html>>.
- eXist <<http://exist.sourceforge.net/>>
- Xindice <<http://xml.apache.org/xindice/>> (Apache <<http://www.apache.org/>>)

À noter que les bases de données relationnelles ordinaires peuvent désormais, non seulement stocker du XML (après tout, ce n'est que du texte) mais également effectuer des recherches avec des langages comme XPath. Voici comment faire en PostgreSQL <http://www.throwingbeans.org/postgresql_and_xml_updated.html>. Cela ne transforme pas pour autant le SGBD relationnel en SGBD XML puisque, par exemple, il n'y a aucune indexation des données, obligeant à analyser le document XML à chaque fois.