

Langages de patch pour XML

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 19 Février 2006

<http://www.bortzmeyer.org/langages-patch-xml.html>

Lorsqu'on veut mettre à jour un gros document XML, il serait très intéressant de pouvoir ne transmettre que la différence avec l'ancien document, le patch. Cela permettrait également de mieux documenter la différence entre les deux fichiers. Mais il n'existe pas de langage standard pour exprimer ces patches. Il y a plusieurs propositions concurrentes.

Un exemple tiré d'une des propositions, XML-patch, vaut mieux qu'un long discours :

```
<add sel="doc"><bar id="ert4773">This is a new child</bar></add>
<remove sel="doc/foo[@a='1']"/>
```

Le patch ci-dessus crée un nouvel élément "bar" et détruit un élément "foo". Comme toutes les propositions de langage de patch, celle-ci utilise XPath pour désigner l'endroit où se font les changements.

Les propositions existantes sont :

- REX (<http://www.w3.org/TR/rex/>), qui vient du consortium W3C et qui est bien plus qu'un langage de patch puisqu'il s'agit d'un moyen de décrire des événements DOM distants.
- Xquery Update (<http://www.w3.org/TR/xqupdate/>), également du W3C, et qui est une extension de XQuery, qui avait été prévu pour étendre XPath.
- XUpdate (<http://xmldb-org.sourceforge.net/xupdate/>), de XML :DB, qui semble aujourd'hui abandonné (la dernière version date de 2000).
- DUL (<http://sourceforge.net/projects/diffxml>) (décrit dans l'Internet-draft (<http://www.ietf.org/ID.html>) draft-mouat-xml-patch-00.txt).
- XML-patch (<http://xmlpatch.sourceforge.net/>) (décrit dans le RFC 5261¹). C'est celui que je trouve le plus simple et le plus proche du cahier des charges "un langage de patch". (Avec un autre langage de données, JSON, Rob Sayre a proposé un format de patch (<http://blog.mozilla.com/rob-sayre/2008/02/15/restful-partial-updates/>) sympathique.)

Une fois qu'on aurait un tel langage, il suffirait de modifier un des programmes de calculs de différence entre fichiers XML (comme Python xmldiff (<ftp://ftp.logilab.org/pub/xmldiff>)).

¹Pour voir le RFC de numéro NNN, <http://www.ietf.org/rfc/rfcNNN.txt>, par exemple <http://www.ietf.org/rfc/rfc5261.txt>