



Tutorial Técnico

parte-se de um formato origem (relacional, por exemplo), transforma-se para XML usando uma determinada estrutura, processa-se e depois se armazena o resultado de alguma forma. Na próxima seção, discutiremos aspectos relacionados a esse ponto em mais detalhes.

Armazenando XML

Apesar do volume de pesquisas e esforços que vêm sendo realizados em gerenciamento de dados XML, em sua maioria, as empresas ainda mantêm seus dados armazenados em bancos de dados relacionais, devido principalmente à maturidade da tecnologia de acesso e armazenamento. Abiteboul, Cluet e Milo (1993) foram os primeiros a sugerir que bancos de dados relacionais fossem utilizados para consultar dados de arquivos. Para isso, seria necessário que os arquivos tivessem alguma estrutura intrínseca. Tal requisito remete a documentos XML que possuem um esquema ou DTD associado.

O que se faz então é utilizar bancos de dados relacionais para armazenar documentos XML produzidos em um determinado processamento (um pedido recebido pela empresa, por exemplo). Além disso, dados relacionais já existentes podem ser manipulados em formato XML. Basta utilizar um método para extrair os dados relacionais em formato XML e usá-los onde for necessário (para alimentar um portal, para trocar dados com outra empresa, etc.).

As versões recentes da maioria dos SGBDs de dados comerciais já suportam XML. Oracle, SQL Server, IBM DB2, somente para citar alguns dos maiores SGBDs, suportam tanto o armazenamento de XML no banco de dados quanto a extração de dados relacionais em XML.

Para versões mais antigas, é possível usar algumas das soluções apresentadas na literatura. Citamos aqui apenas alguns trabalhos. O leitor interessado pode consultar [4] para uma discussão mais profunda desses assuntos.

Armazenamento de dados XML em DBs relacionais. Existem várias formas de se armazenar dados XML em bancos de dados relacionais. Algumas criam uma tabela para cada tipo de elemento do documento, outras usam uma tabela de “arestas” para armazenar os elementos e ponteiros para o elemento pai. Em [5], é feito um estudo do desempenho de cada um desses tipos de armazenamento.

Atualização de documentos XML armazenados em BDs relacionais. Apesar de ainda não existir uma linguagem de atualização padrão para XML (a XQuery [1] lida somente com consultas), alguns trabalhos já tratam da tradução automática de atualizações sobre o documento XML para o que está armazenado no banco de dados relacional. Veja [4] para mais detalhes.

Consulta a dados relacionais com resultados em XML. Existem várias propostas que permitem consultar bancos de dados relacionais como se eles fossem XML, retornando resultado nesse formato. Os principais são SilkRoute e XPERANTO. A idéia aqui é considerar uma visão XML sobre o banco de dados relacional e permitir que essa visão seja consultada, resultando em documentos XML menores.

Atualização de BDs através de visões XML. Além de consultar, é interessante que se possa também atualizar as visões XML. Em [4], apresentamos uma proposta para resolver esse problema.

Além de SGBDs relacionais, existem outras alternativas. Os SGBDs XML nativos vêm se aprimorando cada vez mais. Muitos já suportam inclusive o uso de índices. Entre eles podemos citar XIndice, Timber, eXist, Tamino.

Perspectivas de Pesquisa

Existem grupos de pesquisa no mundo todo trabalhando com assuntos relacionados a XML: chamadas de serviços Web embutidas em documentos XML (documentos XML ativos); *streamings* de dados; recuperação de informação na Web; identificação de instâncias equivalentes (usando similaridade ou outra técnica); consultas a bases distribuídas; entre vários outros tópicos. Ainda existe muito a ser feito, então, mãos à obra!

Referências

1. W3C. *World Wide Web Consortium*. 1994. Disponível em <http://www.w3.org/>.
2. Peat, B; Webber, D. *Introducing XML/EDI*. 1997; Disponível em <http://www.geocities.com/WallStreet/Floor/5815/start.htm>.
3. Megginson, D. *SAX - Simple API for XML*. 1998; Disponível em <http://www.saxproject.org/>.
4. Braganholo, V., *From XML to Relational View Updates: applying old solutions to solve a new problem*, Tese de Doutorado, Instituto de Informática, UFRGS. 2004.
5. Florescu, D.; Kossmann, D. *A performance evaluation of alternative mapping schemes for storing XML data in a relational database*. 1999. Relatório Técnico 3684. INRIA, França.

* Programa de Engenharia de Sistemas e Computação (COPPE/UFRJ) – vanessa@cos.ufrj.br