Kweelt:
More than just "yet another framework to query XML!"

Arnaud Sahuguet
University of Pennsylvania
sahuguet@saul.cis.upenn.edu

1. INTRODUCTION

Kweelt is an open source Java framework for querying XML based on the Quilt proposal [1], now XQuery [3]. It is intended to provide a reference implementation for the Quilt language but also to offer a framework for all kinds of experiments related to XML including storage, optimization, query language features, etc.

The goal of Kweelt is to provide a modular and extensible framework to query XML, where users can replace and extend components as they see fit. To support this goal, the framework offers the following salient characteristics. In Kweelt, the query engine is totally isolated from the storage module. The former is responsible for parsing, massaging and evaluating queries; the latter is in charge of manipulating XML. The bridging between both is accomplished via various APIs. This is new for XML query engines!

This design permits to use multiple storage modules (XML back-ends), even within the same query. We think this is a realistic assumption because XML documents have all sort of shapes, sizes, durabilities and there is no one-size-fits-all way to manipulate XML. Moreover people who already have their favorite tools are more likely to adopt a new technology if it is non intrusive and compatible with their *legacy* tools.

The internal query engine has been implemented on purpose (see [2] for details) in a naive way, by looking at Quilt in a procedural rather than declarative way: the structure of the query will dictate its evaluation. The query is parsed into a an abstract syntax tree and interpreted as is rather than compiled.

Some important features of the Kweelt framework are:
Kweelt implements a query language for XML that satisfies all the requirements from the W3C query-language-requirements and run the corresponding use-cases [4].
It supports multiple XML back-ends. Kweelt can query documents stored in ASCII files, binary files, relational databases or any back-end that supports a DOM-like interface. Kweelt provides some built-in node factories for DOM and a generic relational encoding of XML implemented on top of MySQL and DB2. The user can also easily provide her own.
Kweelt is written in Java 1.2, can run on any Java platform and has a small footprint of less than 200kb.
It is extensible. The user can create his/her own user-defined functions (using Java) and make them available inside the query. Kweelt provides various template classes to make the creation of such functions very easy.
Kweelt also comes with the Kweelt Server Pages (KSP) extension, a built-in Cocoon processor. KSP allows to embed Kweelt queries inside any XML page serviced by Cocoon.
Finally, Kweelt is open-source, which makes it a good candidate for educational or research projects.

ACM SIGMOD 2001 May 21-24, Santa Barbara, California, USA

2. DESCRIPTION OF THE DEMO

In our demo, we will showcase the latest version of the Kweelt framework; we will illustrate the use of multiple back-ends for various queries on local and remote documents; and we will show how Kweelt can be used for information integration and XML publishing (via KSP).

We will also demonstrate how to use Kweelt "to instrument" XML experiments related to performance of DOM implementations, XPath optimization, document caching strategies, etc.

The Kweelt framework is available at
`http://db.cis.upenn.edu/Kweelt`.

3. REFERENCES

[1] D. Chamberlin, J. Robie, and D. Florescu. Quilt: An XML Query Language for Heterogeneous Data Sources. In *Invited paper, WebDB-2000*, May 2000.

[2] A. Sahuguet. Kweelt, the Making-of: Mistakes Made and Lessons Learned. Technical Report MS-CIS-00-23, Department of Computer and Information Science, University of Pennsylvania, November 2000. Available at `http://db.cis.upenn.edu/DL/kweelt-TR.pdf`.

[3] W3C. XQuery: A Query Language for XML. W3C Working Draft 15 February 2001. Available from `http://www.w3.org/TR/2001/WD-xquery-20010215/`.

[4] W3C XML Query Working Group. XML Query Requirements, August 2000. `http://www.w3.org/TR/2000/WD-xmlquery-req-20000815`.