CHAPTER 1

■ ■ ■ ■

A Quick Look at Berkeley DB XML

Most developers, especially Unix programmers, are familiar with Berkeley DB (BDB). The embedded database has been an integral part of BSD-based distributions since 1992, which now include Linux and Apple OS X. Core open source projects such as sendmail, Subversion, MySQL, and OpenLDAP add valuable services atop BDB's key/value storage. Sleepycat—the company that owns, develops, and supports BDB—claims an installation base of more than two million. Google, Amazon.com, AOL, Cisco, Motorola, Sun, and HP are all companies that depend on the database as part of critical applications. In short, BDB is about as ubiquitous as software gets.

■**Note** In February 2006 Oracle acquired Sleepycat Software, pulling the most widely used open source database into its product offering. Oracle plans to continue development of Sleepycat's product line and support of its large customer base.

Because it wanted to move into the XML application space, Sleepycat (with the primary participation of John Merrells) developed BDB XML as a layer atop BDB. Today, BDB XML boasts a sophisticated query engine using XQuery with query plan optimization and flexible indexing. It also inherits the transaction features of BDB.

This chapter gives a brief overview of BDB XML for those familiar with the core concepts: embedded databases, XML, and XQuery. Later chapters examine these topics in more depth. The examples in this chapter make use of the BDB XML shell utility, but can be written in any of the programming language supported by BDB XML—including C++, Perl, Python, Java, and PHP, all covered later in this book. (Tcl is also supported in the BDB XML distribution, but is not covered here.)

A Complete Example

For an illustrative example of exactly what BDB XML does, imagine that we have a collection of XML files for books that we intend to sell. A sample is shown in Listing 1-1.

Listing 1-1. *A Sample Book XML File,* 0553211757.xml

```
<book isbn="0553211757">
  <title>Crime and Punishment</title>
  <author id="923117"/>
  <publisher>Bantam Classics</publisher>
  <weight>9.3</weight>
  <pages>576</pages>
</book>
```