Meet the Experts: Why IBM is open sourcing Cloudscape as Derby

IBM Vice President Jon Prial reveals why

Jon Prial

October 14, 2004

At LinuxWorld in San Francisco, IBM® announced that it was releasing Cloudscape™, a fully relational Java™ database as open source to the Apache Software Foundation, which has established it as an incubator project under the name "Derby." This is the latest in a long series of commitments that IBM has made to the open source community. To find out more about the database, and how it fits into both IBM's strategy and the Java application stack, we talked to Jon Prial, IBM Vice President, Marketing, Information Management Software. Here's what he had to say.

Q: At LinuxWorld, IBM announced that it is contributing Cloudscape to the Apache Software Foundation. What is Cloudscape?

A: Cloudscape is a Java-based, fully transactional relational database technology. It's a purely embedded database that can either be used within an application, or as the database for a more traditional client-server application. It has a very small footprint, and there's no need for a database administrator; you just write the application. You make your database calls, and Cloudscape is there for you.

Q: What's the difference between the open source and IBM commercial version?

A: We've contributed the Cloudscape code to the Apache Software Foundation [ASF]. The project at Apache is called "Derby" and future commercial versions of IBM Cloudscape will be based on the Derby code base.

Derby is an incubator project, the usual way external projects become a part of the ASF. But developers can contribute to and download the code under the Apache license now, and we are actively working with our partners and Apache to grow and build the community working on Derby.

For our Cloudscape product, IBM will take snapshots of the Derby code and provide full support as the IBM Cloudscape offering. You can get a free download of the IBM Cloudscape code as well,
and you can purchase support for it from IBM, but we do not intend to charge a license fee for Cloudscape.

**Q: Why did IBM decide to open source Cloudscape?**

**A:** By open sourcing Cloudscape, IBM hopes to accelerate development of Java-based applications and drive more innovation around Linux and Java. So expanding this market expands the market for high-value IBM middleware, hardware, and services. We think it will especially create new business opportunities in areas such as embedded database applications, small business solutions, and Java and Web-based applications.

Developers recognize good code. When we acquired Cloudscape code along with Informix®, our internal development community started talking about it and working with it. So much so, in fact, that today there are already 70 projects within IBM -- many shipping to marketplace, such as WebSphere® Application Server, WebSphere Portal, and IBM Workplace -- using Cloudscape within their offerings. So we knew we had an interesting technology that could become part of the Java stack. We saw a linkage between the Java database and the Java work being done by Apache, and came to realize this is a tremendous project to open source.

Now, IBM has a long-standing commitment to the open source community. Part of the reason that we've been able to build that relationship is because we know that you don't open source code just for the sake of open sourcing code, or to make things easier for yourself.

IBM understands that it must find the right way to build a community and encourage innovation, and we have done this in a number of different ways, such as supporting the creation of Eclipse.org. By working with Apache, we can build a community of developers who use Derby and also help to make it even better. In fact, we're hoping to see the Derby database become as ubiquitous as the Apache HTTP server.

**Q: What sort of things can you do with Cloudscape?**

**A:** One of the most interesting uses for Cloudscape is as a Java data store, tying the database directly to the Java application. As I mentioned, the Cloudscape database is very small; it's just a 2-MB .jar file. And with a Java database, when a Java application wants to use it, there’s no work to be done at all. No administration. It just makes the JDBC calls and accesses the data, and it does that because it happens to be sitting out there as a .jar file.

Cloudscape technology is actually very versatile -- I think it's important for people to recognize that other types of applications can be used here as well. I mentioned that IBM Workplace uses Cloudscape on the client side for automatic replication back to the server. We provided a technology preview of the "IBM Workplace Data Access" scriptable desktop database -- a business user tool for creating client side databases, queries, and reports -- at LinuxWorld.

And people can use the Cloudscape network server feature to access it as a server-based data store. So Cloudscape can also be used as a client-server database, although we expect to see most of the initial usage around Java applications.
Q: What kind of Eclipse tools will IBM be providing?

A: We have plug-ins for Eclipse that will allow you to access the database. In addition, we have tools for people who use Cloudscape and would like to migrate to a more enterprise-class database. They can migrate from Cloudscape to DB2®, for example.

Q: You mentioned "IBM Workplace Data Access." Could you talk about that some more?

A: IBM Workplace Data Access is a tool for business users to create and query database applications that will run on different desktop environments – including both Windows® and Linux. We used the Eclipse framework to develop the GUI and Cloudscape for the underlying database.

We think IBM Workplace Data Access will open up new opportunities all by itself. But it's also a great example of what you can do when you combine a lightweight, cross-platform database like Cloudscape or Derby with a cross-platform client-side framework for application development like Eclipse.

Q: How does Cloudscape fit into the DB2 database portfolio?

A: Cloudscape meets the need of a new and growing market. As a Java-oriented database, it fills a niche -- which complements our other databases very nicely: DB2 Universal Database™, Informix Dynamic Server, DB2 Express, and so on. And because it's built on open standards, any applications written to Cloudscape can easily move to run on DB2 UDB or other databases as needed for more enterprise-class solutions.

Q: How does Derby differ from other open source databases?

A: No other open source database is as well suited as Derby for extending the Java stack like this, and driving innovation through community involvement. IBM is very focused on meeting this need. We've seen other companies toss databases into open source without a real program to build a community. Many of the open source databases have restrictive licenses. Sometimes ISVs actually have to pay for the right to redistribute open source databases. Sometimes they're not fully relational or transactional.

So Derby is unique compared to other open source databases. It's a truly rich functional relational data store. And, it's available without these kinds of commercial licensing restrictions.

Q: Is there anything else you'd like to add?

A: Cloudscape technology is an extension of IBM's larger data product strategy. Our vision is to integrate heterogeneous data stores -- from zSeries® all the way to pervasive devices -- to support structured and unstructured data, from IBM and non-IBM data sources. We think information integration is the key to tying this together, and our DB2 products provide customers with a framework to do this.
A common and a consistent framework for accessing information enables people to do more things -- to leverage data, to gather more insights, and to become more of an on demand business. We see Cloudscape as a key part of that.

© Copyright IBM Corporation 2004
Trademarks
(www.ibm.com/developerworks/ibm/trademarks/)