IBM's Cloud Escapes

James Governor

In a bid to encourage Java adoption and provide its developers with a native database offering, IBM is submitting Cloudscape, its 100% pure Java embedded database, to the Apache Software Foundation for open source stewardship. The project will be called Derby, and the race to lead and consolidate the embedded database market is now on.

Big Blue's database move is bold, if perhaps overdue. Now that Cloudscape has been decoupled from IBM to some extent, it should see wider adoption. Developers and DBAs can add another open source database to an already potent arsenal of options.

Keeping It Lightweight

IBM is not the only major vendor to open source a database: CA, for one, recently announced that it is open sourcing Ingres, a full function, high-end relational database management system (RDBMS). Meanwhile MySQL continues its inexorable rise as a higher end, open source relational platform, growing up fast via technical contributions from the likes of SAP.

"Now that Cloudscape has been decoupled from IBM to some extent, it should see wider adoption."

Cloudscape, however, is a fundamentally different kind of offering. Based on technology acquired from Informix, the lightweight Cloudscape/Derby code base is designed to appeal to developers and ISVs in need of a lightweight, Java based RDBMS to underpin applications such as web sites, point of sale systems, local registries and departmental applications. Its small footprint—about 2Mb--makes it useful in such deployments.

The "keep it lightweight" approach is in contrast to a product like MySQL, which has continued to add features and functions in order to rival enterprise database stalwarts like Oracle and DB2, and take on workloads such as online transaction processing (OLTP) and decision support systems (DSS).

Derby is somewhat more competitive with Sleepycat's Berkeley DB, although that platform has its own niche and is functionally quite distinct. Sleepycat underpins large-scale, web-based Architectures of Participation including Amazon and Google - web services that we're all familiar with - as well as applications like Evolution and Sun's Java Enterprise System. The database in these cases may be invisible, but Berkeley touches tens of millions of end-users every day and they don't know it. That's the very definition of a successfully embedded database.

IBM and the Cloudscape Diffusion

An important facet of the Cloudscape handoff is that it will also almost certainly have a consolidating impact on IBM's broad database portfolio. Momentum will lead to IBM portfolio rationalization because the market, rather than IBM, will be dictating platform adoption.

IBM originally obtained Cloudscape through its acquisition of Informix, which added a number of different code bases to IBM's DB2 portfolio. Within the broad product set, which features DB2 as the primary money-maker, Cloudscape perhaps didn't receive the sales and marketing attention it deserved. IBM's commitment to customer investment protection however meant that all Informix platforms were supported going forward. This commitment makes product portfolio rationalization decisions particularly difficult.

As Derby takes off it will become clearer to IBM which platforms are going to have the most momentum, and thus deserve service, documentation and development resource allocation going forward. Tough decisions about rationalization become easier when the

"Consider how Cloudscape achieved escape velocity: it was adopted virally within IBM Software Group (SWG), rather than by decree."

community makes them for you.

Consider how Cloudscape achieved escape velocity: it was adopted *virally* within IBM Software Group (SWG), rather than by decree. That is, SWG developers adopted Cloudscape of their own accord because of its simplicity. It is lightweight, easy to administer and fits in very elegantly with Java code - deploying the platform is as straightforward as dropping in a JAR file — adoption of the platform was a choice, **not** a mandate.

So what did IBM's developers do with it? Cloudscape underpins IBM's Workplace, providing synchronization and offline storage capabilities for the rich client. It also stores mappings for IBM Directory Server. It is embedded in over 70 different IBM products, from Lotus to WebSphere. This adoption raised Cloudscape's profile sufficiently that SWG could better see its potential from a corporate perspective. Open sourcing was the obvious approach to propagate the platform.

Big Blue hopes the same virtues of the platform that drove internal adoption will subsequently be replicated outside: the decision to open source is obviously not simple altruism. On the contrary Apache's stewardship is crucial for ensuring that the wider community doesn't dismiss the initiative as an IBM only play, which was an accusation sometimes leveled at the open source Eclipse application lifecycle platform. If Derby is widely adopted outside IBM it will benefit the vendor, because it will encourage adoption of a technology that underpins many of SWG's middleware offerings.

There will be potentially less friction between internal IBM groups. Decisions about platform support can be based on adoption, rather than politics. Cloudscape has in some senses escaped DB2 shackles, but it also makes life easier for the IBM Information Management crowd, helping to rationalize a broad product portfolio with some overlaps.

The Apache Rain Dance

IBM would not be able to popularize Cloudscape by itself, no matter how much marketing muscle it has with ISVs. The firm has shown its open source commitment with its Eclipse contribution and stewardship model, but Cloudscape is targeted at the embedded development space, and database choices are as strategic as they come. Few vendors want to rely on Big Blue alone for their runtimes, certainly not those vendors that compete with IBM.

Apache is patently an ideal home for Cloudscape, even though it might have appeared that the Eclipse Foundation was a more likely destination. IBM is in the middle of a charm offensive with Apache, designed to give the foundation more clout as Big Blue considers how to wrest Java from Sun's grip by pushing it into the open source community. Apache is a king maker when it comes to open source Java, and IBM is well aware of this fact. It therefore makes a great deal of sense for IBM to buddy up to Apache, and giving it some code has got to help in that regard. Initially, IBM and Apache will work together to examine Derby through the Apache incubator model, before proceeding with a broader program of work.

Apache is already the de facto industry standard Web Server. So why not add a database to the stack? It's a natural progression. Developers look to Apache for leadership because of its proven track record.

RedMonk Take

Open source choices are always potentially disruptive, and we believe that Derby will have a significant market impact. It is an excellent choice for Java developers and need not scare off the anti-IBM crowd. The pressure on Microsoft from Java and open source continues apace.

About RedMonk

RedMonk is a research and advisory services firm that assists enterprises, vendors, systems integrators and corporate finance analysts in the decision making process around today's enterprise software stacks. We cover the industry by looking at integrated software stacks, focusing on business and operational context rather than speeds and feeds and feature tick-lists.

Founded by James Governor and Stephen O'Grady, and headquartered in Bath, Maine, RedMonk is on the web at www.redmonk.com. If you would like to discuss this report email jgovernor@redmonk.com.