

BI is Ready for Open Source - Is Open Source Ready for BI?

Article published in DM Review Magazine
August 2005 Issue

By Sam Mohamad

Thomas Edison introduced the world to electricity in 1879 with great fanfare. In contrast, the world hardly noticed when, more than 25 years later, in 1904, a gentleman named Harvey Hubbell was awarded a government patent for his "separable plug." Hubbell, more interested in practicality than PR, wanted to create a product that would eliminate the need to wire all appliances into a building's central power supply. The patent application stated that Hubbell's plug and corresponding socket would allow "persons having little or no electrical knowledge or skill" to connect the growing number of electrical appliances.¹

With its quiet emergence in the 1980s, the open source movement has followed a similar trajectory, eschewing marketing fanfare in favor of creating a more practical, democratic way to distribute technology. The open source revolution makes technology available to everyone and anyone - enabling the community to benefit from the creativity of the whole, resulting in a high quality product that is "self-policed" by a group of committed contributors.

Much like the advent of the plug in last century, open source is inevitable. It gives control to the customer. Enterprise-Class Open Source

Why enterprise-class open source?

Because it works. Think Linux. Think Apache. Think JBoss. These and other open source projects are transforming enterprise IT by leveraging the participatory nature of open source to deliver products and services with superior capabilities and lower prices.

The open source model creates better software by encouraging collaboration - the best idea wins - not only within a company, but throughout a connected, passionate worldwide community. Users can see the code, change it and learn from it; and bugs are more quickly found and fixed. As a result, the open source model often builds more stable, more secure and more easily integrated software - all at a faster pace and lower cost.

This transparency means that commercial open source providers must consistently serve customers through extraordinary value, performance, ease of integration and management. If they fail to achieve this, companies are free to choose another vendor. With open source, technology lock-in and vendor monopolies could become as outdated as steam engines.

Let's contrast this with the proprietary model where product development occurs within one company. This company charges customers to use its software, charges them to fix bugs and problems, and then demands additional charges for upgrades. Open source frees companies from this cycle and puts the power of choice and innovation back into their hands.

Enterprises Want In

The momentum is building. Open source has moved well beyond operating systems and Web servers to more mission-critical systems and to the enterprise application level. As demonstrated by the increased adoption of open source systems by large companies, government agencies (such as the National Weather Service) and educational institutions, organizations find these systems are often easier to operate and maintain better than the proprietary systems they replaced.

With the wide availability of open source products, support resources and a growing community of developers, we see companies beginning to rethink future IT strategies - implementing software that works more easily with users and benefiting from a large, committed community of programmers and businesses.

Business Intelligence is a Market Primed for Open Source

As companies invest to stay compliant with the host of new government regulations, improve operational efficiency and retain customers, business intelligence (BI) has continued its march up the corporate food chain. BI is a top IT priority. At \$16.8 billion in 2001 and growing to \$29 billion in 2006, the BI and data warehousing market is exploding.²

The pain caused by regulatory compliance requirements is exacerbated by the skyrocketing amount of data bombarding companies due to the growth of Internet-based services, wireless communication and personalization.

Despite its importance and executive-level visibility, the database infrastructure on which BI systems are built has evolved very little within the last 20 years. As a result, BI remains expensive, timeconsuming to implement and upgrade, and confined to a limited set of applications. Database software, full-featured reporting, ETL (extract, transform and load) software and other essential pieces of the BI puzzle still carry hefty price tags and require significant time and expertise to deploy.

Most BI implementations fall under the domain of specialized subsections of enterprise IT departments. In the end, most companies pay well more than \$1 million per terabyte of data for BI.

The BI industry's substantial size, prohibitively high costs, and technological and social foundations make it one of the most logical targets for open source. Companies such as JasperSoft, Greenplum and other startups recognize this and are working to provide the focus and funding needed to bring the transformative power of open source to the business intelligence industry. Recent open source product offerings from more traditional BI leaders such as Business Objects and Cognos reinforce the growing importance of the technology for BI.

Is Open Source Ready for BI?

While open source databases and applications have gained a foothold in the enterprise BI market, they have much to prove. However, for some components of BI, early technological foundations exist today, in open source form, with established open source communities.

The existence of these foundations and their active developer communities simplifies, and will accelerate, the route to a BI-focused, open source technology stack.

At a minimum, the new open source products in both data warehousing and BI applications will create new price and technical competition that can only benefit software developers and end users.

Delivering the Promises

Open source databases continue to add the features and customer service elements that are critical for enterprise support. One example of this is the PostgreSQL relational database system (RDBMS).

PostgreSQL hails from the POSTGRES project at the University of California at Berkeley. Professor Michael Stonebraker started the project in 1986 to replace the aging Ingres RDBMS; and DARPA, the National Science Foundation, the Army Research Office and ESL, Inc. sponsored it. While known as the POSTGRES project, the database assumed various roles in different organizations, including an asteroid tracking database, a financial data analysis system and an educational tool.

POSTGRES originally used a language called PostQUEL for accessing database information. In 1994, Andrew Yu and Jolly Chen added the POSTGRES SQL interpreter, originally known as Postgres95.

Postgres95 was then re-licensed under the Berkley software license and shortly thereafter was renamed PostgreSQL.

Along with Ingres, PostgreSQL shares a technical lineage with Sybase, SQL Server and Informix, and is widely considered the most robust of the open source databases.

Today, most people encounter PostgreSQL without even realizing it: Afilias, which manages the .ORG registration, uses PostgreSQL to store all of the .ORG registry information. BASF uses PostgreSQL in a shopping platform for its agriculture products. The World, a media company, has built much of its infrastructure around the use of PostgreSQL.

Simple Licensing

In addition to its strong heritage in data heavy industries, PostgreSQL has a much simpler licensing scheme than other open source alternatives. It is released under the Berkley License, which allows for any use as long as a copy of the Berkley License is included with it. This means that you can release a commercial product that uses PostgreSQL or is a derivative of PostgreSQL without including source code.

Robust Features

PostgreSQL provides more features than alternatives such as MySQL. These include more SQL functions, server-side procedural languages and sophisticated methods for data manipulation. PostgreSQL also offers object-relational capabilities and geometric data types. For users developing an application that has highly complex business rules such as BI, PostgreSQL lets them handle business logic on the database server.

A good way to differentiate databases and test overall quality is to perform an ACID test. ACID is an acronym that describes four properties of a robust database system: atomicity, consistency, isolation and durability. Unlike other open source databases, PostgreSQL is ACID compliant and is extremely responsive in high volume environments, an essential feature for data warehousing and BI deployments.

Strong Support

Unlike other popular open source alternatives, the PostgreSQL community is a true open source community not controlled by a single commercial entity. It is led by an independent steering committee and involves hundreds of developers around the world. This creates a larger pool of support.

PostgreSQL has a reputation for reliability - it is extremely common for companies to report that PostgreSQL has never crashed for them in several years of high activity operation. Not even once. It just works.

But Can It Do BI?

Despite all its benefits, in its current form PostgreSQL lacks many of the features that will enable it to effectively compete with competitive BI solutions. There are several projects afoot in the PostgreSQL community, such as the Greenplum-sponsored Bizgres (www.bizgres.org) project, that are aimed at making open source more relevant to BI by building a complete database system for business intelligence exclusively from free software. The Bizgres Project is a Greenplum sponsored and community supported open source project. According to the PostgreSQL community, most deployments will likely be for departmental applications, though enterprise workloads will also be common.

These and other open source projects, such as those begun by recent open source entrant, EnterpriseDB, will help to make PostgreSQL a powerful, cost effective, and highly supported alternative to Oracle, Sybase, Informix and Microsoft proprietary databases.

Role in BI

It is no longer a question of whether or not open source will be an important enterprise technology, but rather what kind of role it will play. Because of its high price point, long deployment cycles and reliance on specialized expertise, BI is an area that could clearly benefit from alternative approaches.

Never has there been a better time for companies to look at open source as an option for their BI solutions. Innovative startups such as Greenplum, JasperSoft and EnterpriseDB are working to increase the accountability of open source databases by making the features and functionality more relevant for enterprise applications and by collaborating to provide support, more comprehensive and up-to-date documentation, and extreme reliability.

It remains to be seen whether or not business intelligence and data warehousing will be one of the key enterprise markets for open source. However, as open source accountability continues to increase and business users become more and more comfortable with open source, it is likely that less expensive, more modular open source approaches will take hold.

References:

1. National Geographic, June 2005.
2. Gartner.

VISIONTEN

Sam Mohamad brings decades of leadership experience to Greenplum, including five years in senior management positions at Exodus, where he led the company through 18 consecutive quarters of more than 20 percent growth. He spent 10 years leading sales and marketing at Oracle Corporation, where he started and built the company's worldwide service sales organization. Mohamad also led the direct sales and marketing division, where he more than doubled the size of Oracle's telesales business in one year. He previously served as vice president of marketing and strategic sales at Genuity and held sales management positions at Borland and Eastman Kodak.

Reprinted in entirety by VisionTen.

Full credits and copyright is credited to Sam Mohamad and DM Review magazine.
(www.dmreview.com)

The original article can be found [here](#).

http://www.dmreview.com/article_sub.cfm?articleId=1033578

VisionTen is a global IT service and development company, providing open source enterprise class systems for BI, CRM, CMS, ECM, ERP, SCM and other custom business solution.

www.visionten.net