

Chapter 3.1

Open Source Database Technologies

Emmanuel Udoh
Purdue University, USA

INTRODUCTION

The free or open source software (OSS) movement, pioneered by Richard Stallman in 1983, is gaining mainstream acceptance and challenging the established order of the commercial software world. The movement is taking root in various aspects of software development, namely operating systems (Linux), Web servers (Apache), databases (MySQL), and scripting languages (PHP) to mention but a few. The basic tenet of the movement is that the underlying code of any open source software should be freely viewable, modifiable, or redistributable by any interested party, as enunciated under the copyleft concept (Stallman, 2002) This is in sharp contrast to the proprietary software (closed source), in which the code is controlled under the copyright laws.

In the contemporary software landscape, the open source movement can no longer be overlooked by any major players in the indus-

try, as the movement portends a paradigm shift and is forcing a major rethinking of strategy in the software business. For instance, companies like Oracle, Microsoft, and IBM now offer the lightweight versions of their proprietary flagship products to small—to-medium businesses at no cost for product trial (Samuelson, 2006). These developments are signs of the success of the OSS movement. Reasons abound for the success of the OSS, viz. the collective effort of many volunteer programmers, flexible and quick release rate, code availability, and security. On the other hand, one of the main disadvantages of OSS is the limited technical support, as it may be difficult to find an expert to help an organization with system setup or maintenance. Due to the extensive nature of OSS, this article will only focus on the database aspects.

A database is one of the critical components of the application stack for an organization or a business. Increasingly, open-source databases

(OSDBs) such as MySQL, PostgreSQL, MaxDB, Firebird, and Ingress are coming up against the big three commercial proprietary databases: Oracle, SQL server, and IBM DB (McKendrick, 2006; Paulson, 2004; Shankland, 2004). Big companies like Yahoo and Dell are now embracing OSDBs for enterprise-wide applications. According to the Independent Oracle Users Group (IOUG) survey, 37% of enterprise database sites are running at least one of the major brands of open source databases (McKendrick, 2006). The survey further finds that the OSDBs are mostly used for single function systems, followed by custom home-grown applications and Web sites. But critics maintain that these OSDBs are used for nonmission critical purposes, because IT organizations still have concerns about support, security, and management tools (Harris, 2004; Zhao & Elbaum, 2003).

Undoubtedly, the OSDB initiative plays a major role in the IT world, but the expressed concerns about its adoption are the internal IT operations issue of every company. Some companies have successfully integrated OSDBs by using in-house expertise and support. It is therefore incumbent on every company to determine the cost-effectiveness of OSDB adoption before embracing such systems. Furthermore, MySQL database is currently equipped with several features that facilitate the integration with other information systems such as legacy systems and existing software applications. Notably, MySQL converts and imports other databases using a migration toolkit or workbench. This is a powerful framework that supports the migration of several systems with proven methodology.

However, open source databases are forging ahead with attractive business strategies such as the blending of the roles of a database administrator and a developer. Currently, two business models are evolving in the OSDBs space, namely the MySQL and PostgreSQL models (Maguire, 2003; McKendrick, 2006). The MySQL model is based on a dual-licensing approach, whereby a

single firm releases, maintains, and supports the code, while the PostgreSQL model has a community (regulated by experts) entrusted with code release and maintenance. Furthermore, there are more than 30 approved open source licenses, with GNU General Public License (GPL) being the most common (Rosen, 2004). The GNU GPL is copylefted, in that any redistributor of a free software cannot restrict the redistribution or modification of that software. For a detailed treatise on OSS licensing, interested readers are referred to the open source initiative Web site. Due to the large number of OSDBs in the market, this article will further shrink the focus to only the MySQL database and its use with PHP in generating dynamic Web content.

BACKGROUND

The history of MySQL started about three decades ago, when a Swedish firm TcX launched a screen builder-reporting application. Currently, MySQL dominates the OSDBs market with over 6 million installations and with characteristics such as easy administration, performance, stability, robustness, and compactness (LeClaire, 2006; McKendrick, 2006). To underline the popularity of the MySQL database, companies like Dell now package it for distribution with other components of the LAMP (Linux, Apache, MySQL, PHP/Python/Perl) stack. With significant cost savings, businesses deploy it to power high-volume Web sites, critical enterprise applications, and other software packages. A well documented case of MySQL use is Yahoo Finance (Zawodny, 2002). According to Zawodny (2002), MySQL database is cost-effective, easy-to-use, and reliable. It handled high-volume 260 million record tables just as well as the small, low-volume ones. The report further noted key MySQL features such as replication strength, and the ready-to-use APIs and libraries that have helped Yahoo manage its demanding applications smoothly.

5 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/open-source-database-technologies/7947

Related Content

Applying UML for Designing Multidimensional Databases and OLAP Applications

Juan Trujillo, Sergio Lujan-Mora and Il-Yeol Song (2003). *Advanced Topics in Database Research, Volume 2* (pp. 13-36).

www.irma-international.org/chapter/applying-uml-designing-multidimensional-databases/4339

Oceanographic Data Management: Quills and Free Text to the Digital Age and “Big Data”

Justin J. H. Buck and Roy K. Lowry (2017). *Oceanographic and Marine Cross-Domain Data Management for Sustainable Development* (pp. 1-22).

www.irma-international.org/chapter/oceanographic-data-management/166834

Long-Term Fine-Grained Forecasts of Emergency Demand Using EMS Big Data and Population Estimates

Masaki Kaneda, Masahide Nakamura, Sinan Chen and Sachio Saiki (2024). *International Journal of Big Data Intelligence and Applications* (pp. 1-15).

www.irma-international.org/article/long-term-fine-grained-forecasts-of-emergency-demand-using-ems-big-data-and-population-estimates/358106

Cost and Service Capability Considerations on the Intention to Adopt Application Service Provision Services

Yurong Yao, Denis M.S. Lee and Yang W. Lee (2012). *Cross-Disciplinary Models and Applications of Database Management: Advancing Approaches* (pp. 298-322).

www.irma-international.org/chapter/cost-service-capability-considerations-intention/63671

Semantic Enrichment in Knowledge Repositories: Annotating Semantic Relationships Between Discussion Documents

Chih-Ping Wei, Tsang-Hsiang Cheng and Yi-Chung Pai (2006). *Journal of Database Management* (pp. 49-66).

www.irma-international.org/article/semantic-enrichment-knowledge-repositories/3347