

THE EXPERT'S VOICE® IN OPEN SOURCE

Beginning MySQL Database Design and Optimization

From Novice to Professional

*Maximize MySQL database performance by mastering
table design, query, and API best practices*

Jon Stephens
and Chad Russell

APress Media, LLC

Praise for *Beginning MySQL Database Design and Optimization: From Novice to Professional*:

“Beginning MySQL Database Design and Optimization is a great book for MySQL users who already know the basics and want to improve their use of MySQL.”

—Mike Hillyer, Mike Hillyer’s Personal Web Space (www.openwin.org/mike)

“This is a well-rounded volume on MySQL design. There are excellent examples, and the flow of the text is conversational without being rambling and unstructured. The authors have obviously taken great pains to minimize tangents and extraneous information; pithy, but with sufficient detail in mind. The reader is left with neither the sense of being overwhelmed nor longing for an explanation for a glossed-over topic. This book is pretty much a ‘must-have’ for a MySQL programmer looking to bridge the gap between novice and professional.”

—Mary Norbury-Glaser, Slashdot contributor (www.slashdot.org)

“This is a fine book for anyone who works with MySQL.”

—Jack Herrington, Code Generation Network (www.codegeneration.net)

Beginning MySQL Database Design and Optimization: From Novice to Professional

JON STEPHENS AND CHAD RUSSELL

Beginning MySQL Database Design and Optimization: From Novice to Professional
Copyright © 2004 by Jon Stephens and Chad Russell

Originally published by Apress® in 2004

All rights reserved. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage or retrieval system, without the prior written permission of the copyright owner and the publisher.

ISBN 978-1-59059-332-5 ISBN 978-1-4302-0729-0 (eBook)
DOI 10.1007/978-1-4302-0729-0

Trademarked names may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, we use the names only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

Lead Editors: Dominic Shakeshaft and Jason Gilmore

Technical Reviewer: Mike Hillyer

Editorial Board: Steve Anglin, Dan Appleman, Ewan Buckingham, Gary Cornell, Tony Davis,
Jason Gilmore, Chris Mills, Dominic Shakeshaft, Jim Sumser

Project Manager: Tracy Brown Collins

Copy Edit Manager: Nicole LeClerc

Copy Editors: Ami Knox and Marilyn Smith

Production Manager: Kari Brooks-Copony

Production Editor: Katie Stence

Compositor: Dina Quan

Proofreader: Christy Wagner

Indexer: Kevin Broccoli

Artist: Kinetic Publishing Services, LLC

Cover Designer: Kurt Krames

Manufacturing Manager: Tom Debolski

The information in this book is distributed on an “as is” basis, without warranty. Although every precaution has been taken in the preparation of this work, neither the author(s) nor Apress shall have any liability to any person or entity with respect to any loss or damage caused or alleged to be caused directly or indirectly by the information contained in this work.

Contents at a Glance

<i>About the Authors</i>	<i>.ix</i>
<i>About the Technical Reviewer</i>	<i>.xi</i>
<i>Acknowledgments</i>	<i>.xiii</i>
<i>Introduction</i>	<i>.xv</i>
Chapter 1 <i>Review of MySQL Basics</i>	<i>.1</i>
Chapter 2 <i>MySQL Column and Table Types</i>	<i>.45</i>
Chapter 3 <i>Keys, Indexes, and Normalization</i>	<i>.113</i>
Chapter 4 <i>Optimizing Queries with Operators, Branching, and Functions</i>	<i>.171</i>
Chapter 5 <i>Joins, Temporary Tables, and Transactions</i>	<i>.239</i>
Chapter 6 <i>Finding the Bottlenecks</i>	<i>.273</i>
Chapter 7 <i>MySQL Programming</i>	<i>.315</i>
Chapter 8 <i>Looking Ahead</i>	<i>.417</i>
Index	<i>.489</i>

Contents

<i>About the Authors</i>	<i>ix</i>
<i>About the Technical Reviewer</i>	<i>xi</i>
<i>Acknowledgments</i>	<i>xiii</i>
<i>Introduction</i>	<i>xv</i>
Chapter 1 Review of MySQL Basics	1
<i>How to Connect to MySQL</i>	<i>2</i>
<i>Identifiers and Naming Conventions</i>	<i>5</i>
<i>Queries Review</i>	<i>14</i>
<i>Summary</i>	<i>43</i>
<i>What's Next</i>	<i>44</i>
Chapter 2 MySQL Column and Table Types	45
<i>Why Datatypes Matter</i>	<i>46</i>
<i>MySQL Column Types</i>	<i>48</i>
<i>MySQL Table Types</i>	<i>103</i>
<i>Summary</i>	<i>110</i>
<i>What's Next</i>	<i>112</i>
Chapter 3 Keys, Indexes, and Normalization	113
<i>Beyond the Spreadsheet Syndrome</i>	<i>114</i>
<i>Rules for Relational Databases</i>	<i>117</i>
<i>Normalization and Data Modeling</i>	<i>121</i>
<i>Keys, Indexes, and Constraints</i>	<i>134</i>
<i>Common Problems and Errors</i>	<i>164</i>
<i>Summary</i>	<i>169</i>
<i>What's Next</i>	<i>170</i>

Chapter 4 Optimizing Queries with Operators, Branching, and Functions	171
Replacing Program Logic with SQL Logic: A Demonstration	172
MySQL Operators	178
MySQL Functions	193
Branching: Making Choices in Queries	231
Our Demonstration Revisited	236
Summary	237
What's Next	237
Chapter 5 Joins, Temporary Tables, and Transactions	239
Joins	240
Temporary Tables	263
Transactions	267
Summary	270
What's Next	271
Chapter 6 Finding the Bottlenecks	273
Configuration Issues	274
Application Logic	305
Summary	312
What's Next	314
Chapter 7 MySQL Programming	315
Overview of MySQL APIs	316
PHP and the mysql Extension	336
PHP 5 and mysqli	351
Perl-DBI	401
Python and MySQLdb	407
Summary	413
What's Next	415

<i>Chapter 8 Looking Ahead</i>	417
<i>MySQL 4.1</i>	417
<i>MySQL 5.0</i>	429
<i>MySQL 5.1</i>	481
<i>Other Expected Improvements</i>	484
<i>Summary</i>	486
<i>Index</i>	489

About the Authors

Jon Stephens has contributed as an author to seven previous books on Web development and related technologies, including *Usable Shopping Carts*, *Professional PHP Web Services*, *Professional JavaScript (Second Edition)*, and *Professional PHP 4 Web Development Solutions*, and has served as a technical reviewer of a dozen or so more on a number of development topics, including PHP, MySQL, XML, JavaScript, and Visual Basic. He was also one of the original developers of phpUDDI, a PHP Web Services library that has since been incorporated into PEAR as PEAR::UDDI. His articles on MySQL, DOM programming, and other topics have appeared in *International PHP* magazine. Jon studied mathematics in university and started his professional programming career in the early 1990s teaching computers how to operate radio stations. Originally from the USA, Jon now resides in Brisbane, Australia, where he works as a PHP developer for Snapsoft Pty Ltd. and lives with his wife, their daughter, and numerous computers and cats. His chief vices are coffee, cigarettes, and cheap paperback novels.



Chad Russell is currently a contract software developer for staffing industry software leader LiquidMedium, LLC and founder of Russell Information Technologies, Inc. (RIT), an enterprise software startup. Chad has worked on numerous enterprise-level projects over the past 5 years, primarily developing and integrating PHP and MySQL-based applications. He is currently busy with RIT developing enterprise-level, cross-platform software solutions and providing IT consulting. Chad, who resides in Jacksonville, Florida, is very active in his church where he has been a member for 23 years. His hobbies include music (playing bass guitar), writing, fishing, hunting, and programming.



About the Technical Reviewer

Mike Hillyer has been using MySQL for more than three years. In that time, he has received both the MySQL Core and MySQL Professional certifications and has spoken at the 2003 and 2004 MySQL User Conferences. Mike is the webmaster of VB/MYSQL.com (<http://www.vbmysql.com>), a site dedicated to helping Visual Basic developers use MySQL, and volunteers as the resident MySQL expert in the Ask the Experts section of SearchDatabase.com (<http://www.searchdatabase.com>). Mike is also the top-ranked MySQL expert at Experts Exchange (<http://www.experts-exchange.com>). In April 2004, Mike joined MySQL AB as a member of the documentation team and now spends his days writing in his basement and trying to take over the world. So far Mike has taken over the basement and is currently battling for the main floor of his house, but his wife seems to be winning.

Acknowledgments

A GREAT MANY PEOPLE HAVE HELPED as MySQL and this book have grown and evolved over the last 10 months, and some of them have labored very hard indeed to bring you a quality addition to your development library.

Dominic Shakeshaft and Jason Gilmore both did an excellent job of technical editing. Dom encouraged me not only to provide good information, but to produce a good narrative as well. Jason worked extensively on Chapters 7 and 8; his close attention to technical matters undoubtedly improved these chapters greatly over the original drafts. Marilyn Smith and Ami Knox, our copy editors, each contributed by exercising a fine eye for detail. Our project manager, Tracy Brown Collins, deserves special mention for keeping it all hanging together, even in the face of widely divergent schedules and time zones. I have grown to admire and respect all of these people for their competence, dedication, and *Just Being Good Folks To Work With*. Also at Apress: Chris Mills, Katie Stence, Tina Nielsen, Doris Wong, and Gary Cornell.

Thanks to Chad Russell for being a good co-author and a great friend and collaborator.

Mike Hillyer served as our technical reviewer and was outstanding in that role. (He went to work for MySQL AB about halfway through the piece—this should tell you something about his caliber.) I would also like to thank the following individuals at MySQL AB for taking the time to respond to my questions and to point me in the direction of the right answers: Zak Greant, Peter Gulutzman, and Oleksandr Byelkin. In addition, Sasha Pachev at MySQL AB reviewed some material for an earlier writing project that Chad and I worked on but which was never published; some of this later found its way into the present book.

I'm also grateful to Scott, Nathanael, Kenny, Mark, and Cyril at Snapsoft, for putting up with my absences from the office (and the fact I sometimes forgot to announce these ahead of time).

I'd also like to mention some people who've provided me with a great deal of moral and other support over these last few months: Gordy in Edinburgh, Luciana Maria in São Paulo, Sam in Indianapolis, Rachel in New Haven, Jody in Phoenix, JMike in Dallas, Carl in Ljusdal, Dawn in Denver, Dilip in Bangalore, Beatriz in Knoxville, Ana in Ciudad Obregón, Arvind in Buffalo, and Irishka in Moscow; everyone at Hiveminds.Info; people who asked good questions or gave good answers at Experts Exchange; the Thomas family; and my mother and father.

Acknowledgments

Portions of Chapters 7 and 8 appeared in a somewhat different form in Issues 05.04, 06.04, and 01.05 of *International PHP* magazine, copyright © 2004 Software & Support Verlag GmbH, Frankfurt/Main, and are reused in this book with the kind permission of the publisher. Our thanks go to Editor-in-Chief Indu Britto for helping to make this possible.

Jon Stephens
Brisbane, 16 September 2004

Many thanks to the following—without their help, this opportunity would not have been possible for me: God, who makes all things possible; and Jon Stephens, for being the greatest co-author to work with and for all his hard work on this project! Special thanks to my darling Kim for putting up with me while I worked to meet deadlines. Thanks to everyone at Apress for such a wonderful opportunity! My key to success—Matthew 6:33.

Chad Russell

Introduction

THE NEED TO ACCESS DATA through an application is a constant, rather than a variable, in programming. It's probably safe to say that nearly any nontrivial application requires some form of storage and retrieval of data, and for any application of significant size and scope, this means using a relational database. MySQL is a very popular choice because it is

- Available free of charge; anyone can download MySQL from <http://www.mysql.com> for the price of the Internet connection.
- Fast, robust, and scalable; while MySQL is optimized for speed, it still can be (and is) used for datastores containing hundreds and even thousands of tables holding millions of records.
- Easy to install on the most common computing platforms, with executable binaries available for Windows, Linux, Solaris, Mac OS, and others. Since the source code is freely available, MySQL can be compiled for platforms for which executables don't exist, as needed.
- Extremely portable between platforms; it's not difficult to move MySQL databases between machines, and MySQL itself behaves very nearly identically on all supported platforms.

However, all too many developers—even seasoned ones—often end up throwing away MySQL's speed and other advantages due to poorly designed, inefficient databases. In many cases, they create extra work for their application (as well as themselves) due to poor choices throughout the scope of the database creation and through the way they perform operations to retrieve their data.

Given MySQL's popularity with web-based applications with PHP, we will be focusing heavily on this area. However, we will also look at other scenarios in which MySQL is employed, and most of our discussion will apply equally well to programming languages other than PHP.