Community Experience Distilled

Apache Solr High Performance

Boost the performance of Solr instances and troubleshoot real-time problems

Surendra Mohan

[PACKT] open source*
community experience distilled
PUBLISHING

Apache Solr High Performance

Boost the performance of Solr instances and
troubleshoot real-time problems

Surendra Mohan

[PACKT] PUBLISHING
open source
community experience distilled

BIRMINGHAM - MUMBAI

Apache Solr High Performance

Every effort has been made in the preparation of this book to ensure the accuracy of the information presented. However, the information contained in this book is sold without warranty, either express or implied. Neither the author nor Packt Publishing, and its dealers and distributors will be held liable for any damages caused or alleged to be caused directly or indirectly by this book.

Packt Publishing has endeavored to provide trademark information about all of the companies and products mentioned in this book by the appropriate use of capitals. However, Packt Publishing cannot guarantee the accuracy of this information.

First published: March 2014

Production Reference: 1180314

Published by Packt Publishing Ltd.
Livery Place
35 Livery Street
Birmingham B3 2PB, UK.

ISBN 978-1-78216-482-1

www.packtpub.com

Cover Image by Glain Clarrie (glen.m.carrie@gmail.com)

Credits

Author
Surendra Mohan

Reviewers
Azaz Desai
Ankit Jain
Mark Kerzner
Ruben Teijeiro

Acquisition Editor
Neha Nagwekar

Content Development Editor
Poonam Jain

Technical Editor
Krishnaveni Haridas

Copy Editors
Mradula Hegde
Alfida Paiva
Adithi Shetty

Project Coordinator
Puja Shukla

Proofreaders
Simran Bhogal
Ameesha Green
Maria Gould

Indexers
Monica Ajmera Mehta
Mariammal Chettiyar

Graphics
Abhinash Sahu

Production Coordinator
Saiprasad Kadam

Cover Work
Saiprasad Kadam

About the Author

Surendra Mohan, who has served a few top-notch software organizations in varied roles, is currently a freelance software consultant. He has been working on various cutting-edge technologies such as Drupal and Moodle for more than nine years. He also delivers technical talks at various community events such as Drupal meet-ups and Drupal camps. To know more about him, his write-ups, and technical blogs, and much more, log on to `http://www.surendramohan.info/`.

He has also authored the book *Administrating Solr*, *Packt Publishing*, and has reviewed other technical books such as *Drupal 7 Multi Sites Configuration* and *Drupal Search Engine Optimization*, *Packt Publishing*, and titles on Drupal commerce and ElasticSearch, Drupal-related video tutorials, a title on Opsview, and many more.

> I would like to thank my family and friends who supported and encouraged me in completing this book on time with good quality.

About the Reviewers

Azaz Desai has more than three years of experience in Mule ESB, jBPM, and Liferay technology. He is responsible for implementing, deploying, integrating, and optimizing services and business processes using ESB and BPM tools. He was a lead writer of *Mule ESB Cookbook*, *Packt Publishing*, and also played a vital role as a trainer on ESB. He currently provides training on Mule ESB to global clients. He has done various integrations of Mule ESB with Liferay, Alfresco, jBPM, and Drools. He was part of a key project on Mule ESB integration as a messaging system. He has worked on various web services and standards and frameworks such as CXF, AXIS, SOAP, and REST.

Ankit Jain holds a bachelor's degree in Computer Science Engineering from RGPV University, Bhopal, India. He has three years of experience in designing and architecting solutions for the Big Data domain and has been involved with several complex engagements. His technical strengths include Hadoop, Storm, S4, HBase, Hive, Sqoop, Flume, ElasticSearch, Machine Learning, Kafka, Spring, Java, and J2EE.

He also shares his thoughts on his personal blog at `http://ankitasblogger.blogspot.in/`. You can follow him on Twitter at `@mynameisanky`. He spends most of his time reading books and playing with different technologies. When not at work, Ankit spends time with his family and friends, watching movies, and playing games.

> I would like to thank my parents and brother for always being there for me.

Mark Kerzner holds degrees in Law, Maths, and Computer Science. He has been designing software for many years and Hadoop-based systems since 2008. He is the President of SHMsoft, a provider of Hadoop applications for various verticals, and a cofounder of the Hadoop Illuminated training and consulting, as well as the coauthor of the *Hadoop Illuminated* open source book. He has authored and coauthored several books and patents.

> I would like to acknowledge the help of my colleagues, in particular Sujee Maniyam, and last but not least, my multitalented family.

Ruben Teijeiro is an experienced frontend and backend web developer who had worked with several PHP frameworks for over a decade. His expertise is focused now on Drupal, with which he had collaborated in the development of several projects for some important organizations such as UNICEF and Telefonica in Spain and Ericsson in Sweden.

As an active member of the Drupal community, you can find him contributing to Drupal core, helping and mentoring other contributors, and speaking at Drupal events around the world. He also loves to share all that he has learned by writing in his blog, `http://drewpull.com`.

> I would like to thank my parents for supporting me since I had my first computer when I was eight years old, and letting me dive into the computer world. I would also like to thank my fiancée, Ana, for her patience while I'm geeking around the world.

Table of Contents