

Pro Spring Batch



Michael T. Minella

Apress®

Pro Spring Batch

Copyright © 2011 by Michael T. Minella

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 978-1-4302-3452-4

ISBN 978-1-4302-3453-1 (eBook)

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

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

President and Publisher: Paul Manning

Lead Editor: Steve Anglin

Technical Reviewer: Manuel Jordan

Editorial Board: Steve Anglin, Mark Beckner, Ewan Buckingham, Gary Cornell, Jonathan Gennick, Jonathan Hassell, Michelle Lowman, James Markham, Matthew Moodie, Jeff Olson, Jeffrey Pepper, Frank Pohlmann, Douglas Pundick, Ben Renow-Clarke, Dominic Shakeshaft, Matt Wade, Tom Welsh

Coordinating Editor: Anita Castro

Copy Editor: Tiffany Taylor and Mary Behr

Compositor: Bytheway Publishing Services

Indexer: BIM Indexing & Proofreading Services

Artist: SPI Global

Cover Designer: Anna Ishchenko

Distributed to the book trade worldwide by Springer Science+Business Media, LLC., 233 Spring Street, 6th Floor, New York, NY 10013. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com.

For information on translations, please e-mail rights@apress.com, or visit www.apress.com.

Apress and friends of ED books may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Special Bulk Sales—eBook Licensing web page at www.apress.com/bulk-sales.

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.

The source code for this book is available to readers at www.apress.com. You will need to answer questions pertaining to this book in order to successfully download the code.

To my best friend and wife, Erica.

Contents at a Glance

■ About the Author	xii
■ About the Technical Reviewer	xiii
■ Acknowledgments	xiv
■ Chapter 1: Batch and Spring	1
■ Chapter 2: Spring Batch 101	11
■ Chapter 3: Sample Job	29
■ Chapter 4: Understanding Jobs and Steps	45
■ Chapter 5: Job Repository and Metadata	99
■ Chapter 6: Running a Job	119
■ Chapter 7: Readers	165
■ Chapter 8: Item Processors	233
■ Chapter 9: Item Writers	259
■ Chapter 10: Sample Application	331
■ Chapter 11: Scaling and Tuning	387
■ Chapter 12: Testing Batch Processes	447
■ Index	471

Contents

■ About the Author	xii
■ About the Technical Reviewer	xiii
■ Acknowledgments	xiv
■ Chapter 1: Batch and Spring	1
A History of Batch Processing	2
Batch Challenges	3
Why Do Batch Processing in Java?	4
Other Uses for Spring Batch	6
The Spring Batch Framework.....	7
Defining Jobs with Spring	8
Managing Jobs	8
Local and Remote Parallelization	9
Standardizing I/O	9
The Spring Batch Admin Project.....	9
And All the Features of Spring.....	10
How This Book Works.....	10
Summary	10
■ Chapter 2: Spring Batch 101	11
The Architecture of Batch.....	11
Examining Jobs and Steps	12
Job Execution	13

Parallelization	14
Batch Administration	16
Documentation	17
Project Setup	19
Obtaining Spring Batch.....	19
It's the Law: Hello, World!.....	20
Running Your Job	22
Exploring the JobRepository	23
Job Repository Configuration	23
The Job Repository Tables	24
Summary	28
■ Chapter 3: Sample Job.....	29
Understanding Agile Development	29
Capturing Requirements with User Stories	30
Capturing Design with Test-Driven Development.....	32
Using a Source-Control System.....	32
Working with a True Development Environment	32
Understanding the Requirements of the Statement Job.....	33
Designing a Batch Job.....	39
Job Description.....	40
Understanding the Data Model	42
Summary	43
■ Chapter 4: Understanding Jobs and Steps.....	45
Introducing a Job.....	45
Tracing a Job's Lifecycle.....	47
Configuring a Job.....	48
Basic Job Configuration.....	49

Job Inheritance.....	51
Job Parameters	52
Working with Job Listeners.....	59
ExecutionContext.....	61
Manipulating the ExecutionContext.....	62
Working with Steps.....	65
Chunk vs. Item Processing	65
Step Configuration.....	66
Understanding the Other Types of Tasklets.....	68
Step Flow.....	83
Item Error Handling	93
Summary	98
■ Chapter 5: Job Repository and Metadata	99
Configuring the Job Repository	99
Using an In-Memory Job Repository	99
Database.....	101
Using Job Metadata.....	104
The JobExplorer.....	104
The JobOperator	108
Summary	118
■ Chapter 6: Running a Job	119
Starting a Job	119
Job Execution	119
Spring Batch Job Runners.....	120
Third-Party Integration	124
Stopping a Job.....	142
The Natural End.....	142

Programmatic Ending	142
External Stoppage	155
Error Handling.....	157
Controlling Restart.....	160
Preventing a Job from Being Rerun.....	160
Configuring the Number of Restarts	161
Rerunning a Complete Step.....	162
Summary	163
■ Chapter 7: Readers	165
The ItemReader Interface	165
File Input.....	166
Flat Files	166
XML.....	198
Database Input.....	201
JDBC.....	202
Hibernate	208
JPA	214
Existing Services	218
Custom Input	221
Error Handling.....	226
Skipping Records.....	226
Logging Invalid Records	228
Dealing with No Input	230
Summary	231
■ Chapter 8: Item Processors	233
Introduction to ItemProcessors.....	233

Using Spring Batch’s ItemProcessors.....	235
ValidatingItemProcessor.....	235
ItemProcessorAdapter.....	243
CompositeltemProcessor.....	248
Writing Your Own ItemProcessor.....	256
Filtering Items.....	257
Summary.....	258
■ Chapter 9: Item Writers.....	259
Introduction to ItemWriters.....	260
File-Based ItemWriters.....	261
FlatFileItemWriter.....	261
StaxEventItemWriter.....	273
Database-Based ItemWriters.....	278
JdbcBatchItemWriter.....	278
HibernateItemWriter.....	285
JpaltelWriter.....	290
Alternative Output Destination ItemWriters.....	294
ItemWriterAdapter.....	294
PropertyExtractingDelegatingItemWriter.....	296
JmsItemWriter.....	298
SimpleMailMessageItemWriter.....	302
Multipart ItemWriters.....	310
MultiResourceItemWriter.....	311
CompositeltemWriter.....	322
ClassifierCompositeltemWriter.....	326
Summary.....	330

■ Chapter 10: Sample Application	331
Reviewing the Statement Job.....	331
Setting Up a New Project.....	332
Importing Customer and Transaction Data	336
Creating the Customer Transaction Reader.....	338
Looking Ip Ids	346
Writing the Customer and Transaction Data.....	353
Downloading Current Stock Prices	355
Reading the Tickers.....	356
Writing the Stock File	360
Importing Current Stock Prices.....	361
Reading the Stock Price File.....	362
Writing Stock Prices to the Database	363
Calculating Pricing Tiers	364
Reading How Many Transactions the Customer Had	365
Calculating the Pricing Tier	367
Updating the Database with the Calculated Tier	369
Calculating Transaction Fees.....	370
Reading the Transactions.....	370
Saving Transaction Fees to the Database	374
Generating Monthly Statement.....	375
Reading the Statement Data.....	375
Writing the Statements.....	381
Summary	385
■ Chapter 11: Scaling and Tuning.....	387
Profiling Your Batch Process	387
A Tour of VisualVM.....	388

Profiling Spring Batch Applications	392
Scaling a Job	399
Multithreaded Steps	399
Parallel Steps.....	404
Remote Chunking	421
Partitioning	435
Summary	445
■ Chapter 12: Testing Batch Processes	447
Unit Tests with JUnit and Mockito	447
JUnit	448
Mock Objects.....	451
Mockito.....	452
Integration Tests with Spring Classes	456
General Integration Testing with Spring.....	456
Testing Spring Batch	461
Summary	470
■ Index	471

About the Author



■ **Michael T. Minella** is a software engineer, teacher, and author with over a decade of enterprise development experience using commercial and open source technologies. He is a software engineer for Trustwave and an instructor at DePaul University. He wrote the popular Refcard JUnit and EasyMock, was the technical reviewer for Struts 2 Web Application Development by Dave Newton, and maintains a popular blog on open source technologies.

Michael's technical background runs the gambit. From the first programming language he learned (mainframe Assembler) to the languages he uses and teaches now (Java, JavaScript, Flex/ActionScript, and so on) he has been all over the map. His passions include quality software development, learning new things, and sharing knowledge with others through both mentoring and formal teaching. When he's not in front of a computer, Michael

enjoys spending time with his family and friends as well as pursuing other hobbies including woodworking and reading.

You can follow Michael on Twitter at www.twitter.com/michaelminella or read his blog at www.michaelminella.com.

About the Technical Reviewer



■ **Manuel Jordan Elera** is a freelance Java developer. He designs and develops personal systems for his customers using powerful frameworks based in Java such as Spring and Hibernate, among others. Manuel is now an autodidact developer and enjoys learning new frameworks to get better results in his projects.

Manuel received a degree in systems engineering with honors, and he won the 2010 Springy Award – Community Champion. In his little free time, he likes reading the Bible and composing music with his guitar. Manuel is a senior member of the Spring Community Forums, where his username is dr_pompeii.

Manuel has been the technical reviewer for the follows books: *Pro SpringSource dm Server*, *Spring Enterprise Recipes*, *Spring Recipes* (Second Edition), and *Pro Spring Integration*.

You can contact him through his blog at <http://manueljordan.wordpress.com/>.

Acknowledgments

In the space of a page or so, I can't hope to convey my appreciation to everyone who helped me along the way to making this, my first book, a reality. However, I would like to call out a few specific people who have made extraordinary contributions to the process.

First, anyone who has written a book knows that although there is one name on the cover, it takes one heck of a team to write a book. Thank you to everyone at Apress for giving me the opportunity. Thanks to Steve Anglin for believing in the concept and allowing me the chance to write as a first-time author. Thank you to Anita Castro and Douglas Pundick for your relentless drive to help me finish this book. I'd also like to thank all of the editorial staff who worked on the book; your efforts made me look like I actually knew how to spell and construct an intelligent sentence.

A special thank you to Manuel Jordan, my technical reviewer. Without his thorough feedback on the text and testing of every example, this book would not be nearly what it is now.

Within my professional life, a number of people have helped me along the way to this milestone. First I'd like to thank Peter Costa and the Fees team for giving me the opportunity to work with Spring Batch in such a demanding environment. Thanks to Joel Tosi for his encouragement and for pushing me to expand who I am professionally beyond what happens in a cubicle. I'd also like to thank my manager at Trustwave, Matt Konda: his ongoing support and encouragement is more than I ever expected or could have asked for.

Thanks to David Sayer and the people on the Spring Batch team. You have not only put together a great framework but also built an amazing community around it.

Writing a book takes a toll not only on the author but on everyone in the author's life. Their support allows me to do what I do. I'd like to offer a special thank you to my family and friends who have put time with me on hold to allow me to make this dream come true. Thank you to my parents, my sister, my extended family, my wife's family, and all my friends. All of you sacrificed over the course of the eight months it took me to write this book. I love you all very much.

I conclude with the one person who has supported me without question during this entire process. While she put things on hold for us and made things happen when I couldn't, I wrote my book. My most heartfelt thank you goes to my best friend and wife, Erica. Without you, neither this book nor who I am today would be possible. I love you and thank you.