



Ilya Ganelin  
Ema Orhian  
Kai Sasaki  
Brennon York

# SPARK

---

BIG DATA CLUSTER COMPUTING  
IN PRODUCTION

WILEY



**Spark™**

Big Data Cluster Computing in Production





**Spark™**

---

Big Data Cluster Computing in Production

Ilya Ganelin  
Ema Orhian  
Kai Sasaki  
Brennon York

**WILEY**

## Spark™: Big Data Cluster Computing in Production

Published by  
John Wiley & Sons, Inc.  
10475 Crosspoint Boulevard  
Indianapolis, IN 46256  
www.wiley.com

Copyright © 2016 by John Wiley & Sons, Inc., Indianapolis, Indiana  
Published simultaneously in Canada

ISBN: 978-1-119-25401-0

ISBN: 978-1-119-25404-1 (ebk)

ISBN: 978-1-119-25405-8 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

**Limit of Liability/Disclaimer of Warranty:** The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a particular purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies contained herein may not be suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Neither the publisher nor the author shall be liable for damages arising herefrom. The fact that an organization or Web site is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or website may provide or recommendations it may make. Further, readers should be aware that Internet websites listed in this work may have changed or disappeared between when this work was written and when it is read.

For general information on our other products and services please contact our Customer Care Department within the United States at (877) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at <http://booksupport.wiley.com>. For more information about Wiley products, visit [www.wiley.com](http://www.wiley.com).

**Library of Congress Control Number:** 2016932284

**Trademarks:** Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. Spark is a trademark of The Apache Software Foundation. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.



## About the Authors



**Ilya Ganelin** is a roboticist turned data engineer. After a few years at the University of Michigan building self-discovering robots and another few years working on embedded DSP software with cell phones and radios at Boeing, he landed in the world of Big Data at the Capital One Data Innovation Lab. Ilya is an active contributor to the core components of Apache Spark and a committer to Apache Apex, with the goal of learning what it takes to build a next-generation distributed computing platform. Ilya is an avid bread maker and cook, skier, and race-car driver.



**Ema Orhian** is a passionate Big Data Engineer interested in scaling algorithms. She is actively involved in the Big Data community, organizing and speaking at conferences, and contributing to open source projects. She is the main committer on `jaws-spark-sql-rest`, a data warehouse explorer on top of Spark SQL. Ema has been working on bringing Big Data analytics into healthcare, developing an end-to-end pipeline for computing statistical metrics on top of large datasets.



**Kai Sasaki** is a Japanese software engineer who is interested in distributed computing and machine learning. Although the beginning of his career didn't start with Hadoop or Spark, his original interest toward middleware and fundamental technologies that support a lot of these services and the Internet drives him toward this field. He has been a Spark contributor who develops mainly MLlib and ML libraries. Nowadays, he is trying to research the great potential of combining deep learning and Big Data. He believes that Spark can play a significant role even in artificial intelligence in the Big Data era. GitHub: <https://github.com/Lewuathe>.



**Brennon York** is an aerobatic pilot moonlighting as a computer scientist. His true loves are distributed computing, scalable architectures, and programming languages. He has been a core contributor to Apache Spark since 2014 with the goal of developing a stronger community and inspiring collaboration through development on GraphX and the core build environment. He has had a relationship with Spark since his contributions began and has been taking applications into production with the framework since that time.





## About the Technical Editors

**Ted Yu** is a Staff Engineer at HortonWorks. He is also an HBase PMC and Spark contributor and has been using/contributing to Spark for more than one year.

**Dan Osipov** is a Principal Consultant at Applicative, LLC. He has been working with Spark for the last two years, and has been working in Scala for about four years, primarily with data tools and applications. Previously he was involved in mobile development and content management systems.

**Jeff Thompson** is a neuro-scientist turned data scientist with a PhD from UC Berkeley in vision science (primarily neuroscience and brain imaging), and a post-doc at Boston University's bio-medical imaging center. He has spent a few years working at a homeland security startup as an algorithms engineer building next-gen cargo screening systems. For the last two years he has been a senior data scientist at Bosch, a global engineering and manufacturing company.

**Anant Asthana** is a Big Data consultant and Data Scientist at Pythian. He has a background in device drivers and high availability/critical load database systems.

**Bernardo Palacio Gomez** is a Consulting Member of the Technical Staff at Oracle on the Big Data Cloud Service Team.

**Gaspar Munoz** works for Stratio (<http://www.stratio.com>) as a product architect. Stratio was the first Big Data platform based on Spark, so he has worked with Spark since it was in the incubator. He has put into production several projects

using Spark core, Streaming, and SQL for some of the most important banks in Spain. He has also contributed to Spark and the spark-csv projects.

**Brian Gawalt** received a Ph.D. in electrical engineering from UC Berkeley in 2012. Since then he has been working in Silicon Valley as a data scientist, specializing in machine learning over large datasets.

**Adamos Loizou** is a Java/Scala Developer at OVO Energy.



**Project Editor**

Charlotte Kughen

**Production Editor**

Christine O'Connor

**Technical Editors**

Ted Yu

Dan Osipov

Jeff Thompson

Anant Asthana

Bernardo Palacio Gomez

Gaspar Munoz

Brian Gawalt

Adamos Loizou

**Production Manager**

Kathleen Wisor

**Manager of Content Development  
& Assembly**

Mary Beth Wakefield

**Marketing Director**

David Mayhew

**Marketing Manager**

Carrie Sherrill

**Professional Technology & Strategy  
Director**

Barry Pruett

**Business Manager**

Amy Knies

**Associate Publisher**

Jim Minatel

**Project Coordinator, Cover**

Brent Savage

**Proofreader**

Nancy Carrasco

**Indexer**

John Sleeva

**Cover Designer**

Wiley

**Cover Image**

ktsimage/iStockphoto





# Acknowledgments

We would like to offer a special thank you to Yuichi-Tanaka who worked with Kai to provide the use case example found in Chapter 6.

We would like to acknowledge and thank each of the authors for contributing their knowledge to make this book possible. Further we would like to thank the editors for their time and Wiley as our publisher.

The authors came from various companies and we want to thank the individual companies that were able to aid in the success of this book, even from a secondhand nature, in giving each of them the ability to write about their individual experiences they've had, both personally and in the field. With that, we would like to thank Capital One.

We would also like to thank the various other companies that are contributing in myriad ways to better Apache Spark as a whole. These include, but are certainly not limited to (and we apologize if we missed any), DataBricks, IBM, Cloudera, and TypeSafe.

Finally, this book would not have been possible without the ongoing work of the people who've contributed to the Apache Spark project, including the Spark Committers, the Spark Project Management Committee, and the Apache Software Foundation.





# Contents at a glance

<b>Introduction</b>	<b>xix</b>
<b>Chapter 1 Finishing Your Spark Job</b>	<b>1</b>
<b>Chapter 2 Cluster Management</b>	<b>19</b>
<b>Chapter 3 Performance Tuning</b>	<b>53</b>
<b>Chapter 4 Security</b>	<b>83</b>
<b>Chapter 5 Fault Tolerance or Job Execution</b>	<b>105</b>
<b>Chapter 6 Beyond Spark</b>	<b>145</b>
<b>Index</b>	<b>189</b>







# Contents

<b>Introduction</b>	<b>xix</b>
<b>Chapter 1 Finishing Your Spark Job</b>	<b>1</b>
Installation of the Necessary Components	2
Native Installation Using a Spark Standalone Cluster	3
The History of Distributed Computing That Led to Spark	3
Enter the Cloud	4
Understanding Resource Management	5
Using Various Formats for Storage	8
Text Files	10
Sequence Files	11
Avro Files	11
Parquet Files	12
Making Sense of Monitoring and Instrumentation	13
Spark UI	13
Spark Standalone UI	15
Metrics REST API	16
Metrics System	16
External Monitoring Tools	16
Summary	17
<b>Chapter 2 Cluster Management</b>	<b>19</b>
Background	21
Spark Components	24
Driver	25
Workers and Executors	26
Configuration	27
Spark Standalone	30
Architecture	31
Single-Node Setup Scenario	31

	Multi-Node Setup	32
	YARN	33
	Architecture	35
	Dynamic Resource Allocation	37
	Scenario	39
	Mesos	40
	Setup	41
	Architecture	42
	Dynamic Resource Allocation	44
	Basic Setup Scenario	44
	Comparison	46
	Summary	50
<b>Chapter 3</b>	<b>Performance Tuning</b>	<b>53</b>
	Spark Execution Model	54
	Partitioning	56
	Controlling Parallelism	56
	Partitioners	58
	Shuffling Data	59
	Shuffling and Data Partitioning	61
	Operators and Shuffling	63
	Shuffling Is Not That Bad After All	67
	Serialization	67
	Kryo Registrators	69
	Spark Cache	69
	Spark SQL Cache	73
	Memory Management	73
	Garbage Collection	74
	Shared Variables	75
	Broadcast Variables	76
	Accumulators	78
	Data Locality	81
	Summary	82
<b>Chapter 4</b>	<b>Security</b>	<b>83</b>
	Architecture	84
	Security Manager	84
	Setup Configurations	85
	ACL	86
	Configuration	86
	Job Submission	87
	Web UI	88
	Network Security	95
	Encryption	96
	Event logging	101

	Kerberos	101
	Apache Sentry	102
	Summary	102
<b>Chapter 5</b>	<b>Fault Tolerance or Job Execution</b>	<b>105</b>
	Lifecycle of a Spark Job	106
	Spark Master	107
	Spark Driver	109
	Spark Worker	111
	Job Lifecycle	112
	Job Scheduling	112
	Scheduling within an Application	113
	Scheduling with External Utilities	120
	Fault Tolerance	122
	Internal and External Fault Tolerance	122
	Service Level Agreements (SLAs)	123
	Resilient Distributed Datasets (RDDs)	124
	Batch versus Streaming	130
	Testing Strategies	133
	Recommended Configurations	139
	Summary	142
<b>Chapter 6</b>	<b>Beyond Spark</b>	<b>145</b>
	Data Warehousing	146
	Spark SQL CLI	147
	Thrift JDBC/ODBC Server	147
	Hive on Spark	148
	Machine Learning	150
	DataFrame	150
	MLlib and ML	153
	Mahout on Spark	158
	Hivemall on Spark	160
	External Frameworks	161
	Spark Package	161
	XGBoost	163
	spark-jobserver	164
	Future Works	166
	Integration with the Parameter Server	167
	Deep Learning	175
	Enterprise Usage	182
	Collecting User Activity Log with Spark and Kafka	183
	Real-Time Recommendation with Spark	184
	Real-Time Categorization of Twitter Bots	186
	Summary	186
<b>Index</b>		<b>189</b>





# Introduction

Apache Spark is a distributed compute framework for easy, at-scale, computation. Some refer to it as a “compute grid” or a “compute framework”—these terms are also correct within the underlying premise that Spark makes it easy for developers to gain access and insight into vast quantities of data.

Apache Spark was created by Matei Zaharia as a research project inside of the University of California, Berkeley in 2009. It was donated to the open source community in 2010. In 2013 Spark was added into the Apache Software Foundation as an Incubator project and graduated into a Top Level Project (TLP) in 2014, where it remains today.

## Who This Book Is For

---

If you’ve picked up this book we presume that you already have an extended fascination with Apache Spark. We consider the intended audience for this book to be one of a developer, a project lead for a Spark application, or a system administrator (or DevOps) who needs to prepare to take a developed Spark application into a migratory path for a production workflow.

## What This Book Covers

---

This book covers various methodologies, components, and best practices for developing and maintaining a production-grade Spark application. That said, we presume that you already have an initial or possible application scoped for production as well as a known foundation for Spark basics.

## How This Book Is Structured

---

This book is divided into six chapters, with the aim of imparting readers with the following knowledge:

- A deep understanding of the Spark internals as well as their implication on the production workflow
- A set of guidelines and trade-offs on the various configuration parameters that can be used to tune Spark for high availability and fault tolerance
- A complete picture of a production workflow and the various components necessary to migrate an application into a production workflow

## What You Need to Use This Book

---

You should understand the basics of development and usage atop Apache Spark. This book *will not* be covering introductory material. There are numerous books, forums, and resources available that cover this topic and, as such, we assume all readers have basic Spark knowledge or, if duly lost, will read the interested topics to better understand the material presented in this book.

The source code for the samples is available for download from the Wiley website at: [www.wiley.com/go/sparkbigdataclustercomputing](http://www.wiley.com/go/sparkbigdataclustercomputing).

## Conventions

---

To help you get the most from the text and keep track of what's happening, we've used a number of conventions throughout the book.

**NOTE** Notes indicate notes, tips, hints, tricks, or asides to the current discussion. As for styles in the text:

- We *highlight* new terms and important words when we introduce them.
- We show code within the text like so: `persistence.properties`.

## Source Code

---

As you work through the examples in this book, you may choose either to type in all the code manually, or to use the source code files that accompany the book. All the source code used in this book is available for download at [www.wiley.com](http://www.wiley.com).