

O'REILLY®

第2版



Cassandra

The Definitive Guide

Cassandra权威指南 (影印版)

Jeff Carpenter, Eben Hewitt 著

東南大學出版社

第2版

Cassandra权威指南 (影印版)

Cassandra: The Definitive Guide

Jeff Carpenter, Eben Hewitt 著

Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®

O'Reilly Media, Inc. 授权东南大学出版社出版

南京 东南大学出版社

图书在版编目(CIP)数据

Cassandra 权威指南:第2版:英文/(美)杰夫·卡彭特(Jeff Carpenter),(美)埃本·休伊特(Eben Hewitt)著. —影印本. —南京:东南大学出版社,2018.1

书名原文:Cassandra:The Definitive Guide,2E

ISBN 978-7-5641-7294-7

I. ①C… II. ①杰… ②埃… III. ①关系数据库—数据库管理系统—英文 IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2017)248400 号
图字:10-2017-342 号

© 2016 by O'Reilly Media, Inc.

Reprint of the English Edition, jointly published by O'Reilly Media, Inc. and Southeast University Press, 2018. Authorized reprint of the original English edition, 2017 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由 O'Reilly Media, Inc. 出版 2016。

英文影印版由东南大学出版社出版 2018。此影印版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc. 的许可。

版权所有,未得书面许可,本书的任何部分和全部不得以任何形式重制。

Cassandra 权威指南 第2版(影印版)

出版发行:东南大学出版社

地 址:南京四牌楼2号 邮编:210096

出 版 人:江建中

网 址:<http://www.seupress.com>

电子邮件:press@seupress.com

印 刷:常州市武进第三印刷有限公司

开 本:787毫米×980毫米 16开本

印 张:23.25

字 数:455千字

版 次:2018年1月第1版

印 次:2018年1月第1次印刷

书 号:ISBN 978-7-5641-7294-7

定 价:98.00元

本社图书若有印装质量问题,请直接与营销部联系。电话(传真):025-83791830

*This book is dedicated to my sweetheart, Alison Brown.
I can hear the sound of violins, long before it begins.*

—E.H.

*For Stephanie, my inspiration, unfailing support,
and the love of my life.*

—J.C.

Foreword

Cassandra was open-sourced by Facebook in July 2008. This original version of Cassandra was written primarily by an ex-employee from Amazon and one from Microsoft. It was strongly influenced by Dynamo, Amazon's pioneering distributed key/value database. Cassandra implements a Dynamo-style replication model with no single point of failure, but adds a more powerful "column family" data model.

I became involved in December of that year, when Rackspace asked me to build them a scalable database. This was good timing, because all of today's important open source scalable databases were available for evaluation. Despite initially having only a single major use case, Cassandra's underlying architecture was the strongest, and I directed my efforts toward improving the code and building a community.

Cassandra was accepted into the Apache Incubator, and by the time it graduated in March 2010, it had become a true open source success story, with committers from Rackspace, Digg, Twitter, and other companies that wouldn't have written their own database from scratch, but together built something important.

Today's Cassandra is much more than the early system that powered (and still powers) Facebook's inbox search; it has become "the hands-down winner for transaction processing performance," to quote Tony Bain, with a deserved reputation for reliability and performance at scale.

As Cassandra matured and began attracting more mainstream users, it became clear that there was a need for commercial support; thus, Matt Pfeil and I cofounded Ripitano in April 2010. Helping drive Cassandra adoption has been very rewarding, especially seeing the uses that don't get discussed in public.

Another need has been a book like this one. Like many open source projects, Cassandra's documentation has historically been weak. And even when the documentation ultimately improves, a book-length treatment like this will remain useful.

Thanks to Eben for tackling the difficult task of distilling the art and science of developing against and deploying Cassandra. You, the reader, have the opportunity to learn these new concepts in an organized fashion.

— Jonathan Ellis
*Project Chair, Apache Cassandra, and
Cofounder and CTO, DataStax*

Foreword

I am so excited to be writing the foreword for the new edition of *Cassandra: The Definitive Guide*. Why? Because there is a new edition! When the original version of this book was written, Apache Cassandra was a brand new project. Over the years, so much has changed that users from that time would barely recognize the database today. It's notoriously hard to keep track of fast moving projects like Apache Cassandra, and I'm very thankful to Jeff for taking on this task and communicating the latest to the world.

One of the most important updates to the new edition is the content on modeling your data. I have said this many times in public: a data model can be the difference between a successful Apache Cassandra project and a failed one. A good portion of this book is now devoted to understanding how to do it right. Operations folks, you haven't been left out either. Modern Apache Cassandra includes things such as virtual nodes and many new options to maintain data consistency, which are all explained in the second edition. There's so much ground to cover—it's a good thing you got the definitive guide!

Whatever your focus, you have made a great choice in learning more about Apache Cassandra. There is no better time to add this skill to your toolbox. Or, for experienced users, maintaining your knowledge by keeping current with changes will give you an edge. As recent surveys have shown, Apache Cassandra skills are some of the highest paying and most sought after in the world of application development and infrastructure. This also shows a very clear trend in our industry. When organizations need a highly scaling, always-on, multi-datacenter database, you can't find a better choice than Apache Cassandra. A quick search will yield hundreds of companies that have staked their success on our favorite database. This trust is well founded, as you will see as you read on. As applications are moving to the cloud by default, Cassandra keeps up with dynamic and global data needs. This book will teach you why and how to apply it in your application. Build something amazing and be yet another success story.

And finally, I invite you to join our thriving Apache Cassandra community. Worldwide, the community has been one of the strongest non-technical assets for new users. We are lucky to have a thriving Cassandra community, and collaboration among our members has made Apache Cassandra a stronger database. There are many ways you can participate. You can start with simple things like attending meet-ups or conferences, where you can network with your peers. Eventually you may want to make more involved contributions like writing blog posts or giving presentations, which can add to the group intelligence and help new users following behind you. And, the most critical part of an open source project, make technical contributions. Write some code to fix a bug or add a feature. Submit a bug report or feature request in a JIRA. These contributions are a great measurement of the health and vibrancy of a project. You don't need any special status, just create an account and go! And when you need help, refer back to this book, or reach out to our community. We are here to help you be successful.

Excited yet? Good!

Enough of me talking, it's time for you to turn the page and start learning.

— Patrick McFadin
Chief Evangelist for
Apache Cassandra, DataStax

Preface

Why Apache Cassandra?

Apache Cassandra is a free, open source, distributed data storage system that differs sharply from relational database management systems (RDBMSs).

Cassandra first started as an Incubator project at Apache in January of 2009. Shortly thereafter, the committers, led by Apache Cassandra Project Chair Jonathan Ellis, released version 0.3 of Cassandra, and have steadily made releases ever since. Cassandra is being used in production by some of the biggest companies on the Web, including Facebook, Twitter, and Netflix.

Its popularity is due in large part to the outstanding technical features it provides. It is durable, seamlessly scalable, and tuneably consistent. It performs blazingly fast writes, can store hundreds of terabytes of data, and is decentralized and symmetrical so there's no single point of failure. It is highly available and offers a data model based on the Cassandra Query Language (CQL).

Is This Book for You?

This book is intended for a variety of audiences. It should be useful to you if you are:

- A developer working with large-scale, high-volume applications, such as Web 2.0 social applications or ecommerce sites
- An application architect or data architect who needs to understand the available options for high-performance, decentralized, elastic data stores
- A database administrator or database developer currently working with standard relational database systems who needs to understand how to implement a fault-tolerant, eventually consistent data store

- A manager who wants to understand the advantages (and disadvantages) of Cassandra and related columnar databases to help make decisions about technology strategy
- A student, analyst, or researcher who is designing a project related to Cassandra or other non-relational data store options

This book is a technical guide. In many ways, Cassandra represents a new way of thinking about data. Many developers who gained their professional chops in the last 15–20 years have become well versed in thinking about data in purely relational or object-oriented terms. Cassandra’s data model is very different and can be difficult to wrap your mind around at first, especially for those of us with entrenched ideas about what a database is (and should be).

Using Cassandra does not mean that you have to be a Java developer. However, Cassandra is written in Java, so if you’re going to dive into the source code, a solid understanding of Java is crucial. Although it’s not strictly necessary to know Java, it can help you to better understand exceptions, how to build the source code, and how to use some of the popular clients. Many of the examples in this book are in Java. But because of the interface used to access Cassandra, you can use Cassandra from a wide variety of languages, including C#, Python, node.js, PHP, and Ruby.

Finally, it is assumed that you have a good understanding of how the Web works, can use an integrated development environment (IDE), and are somewhat familiar with the typical concerns of data-driven applications. You might be a well-seasoned developer or administrator but still, on occasion, encounter tools used in the Cassandra world that you’re not familiar with. For example, Apache Ant is used to build Cassandra, and the Cassandra source code is available via Git. In cases where we speculate that you’ll need to do a little setup of your own in order to work with the examples, we try to support that.

What’s in This Book?

This book is designed with the chapters acting, to a reasonable extent, as standalone guides. This is important for a book on Cassandra, which has a variety of audiences and is changing rapidly. To borrow from the software world, the book is designed to be “modular.” If you’re new to Cassandra, it makes sense to read the book in order; if you’ve passed the introductory stages, you will still find value in later chapters, which you can read as standalone guides.

Here is how the book is organized:

Chapter 1, Beyond Relational Databases

This chapter reviews the history of the enormously successful relational database and the recent rise of non-relational database technologies like Cassandra.

Chapter 2, Introducing Cassandra

This chapter introduces Cassandra and discusses what's exciting and different about it, where it came from, and what its advantages are.

Chapter 3, Installing Cassandra

This chapter walks you through installing Cassandra, getting it running, and trying out some of its basic features.

Chapter 4, The Cassandra Query Language

Here we look at Cassandra's data model, highlighting how it differs from the traditional relational model. We also explore how this data model is expressed in the Cassandra Query Language (CQL).

Chapter 5, Data Modeling

This chapter introduces principles and processes for data modeling in Cassandra. We analyze a well-understood domain to produce a working schema.

Chapter 6, The Cassandra Architecture

This chapter helps you understand what happens during read and write operations and how the database accomplishes some of its notable aspects, such as durability and high availability. We go under the hood to understand some of the more complex inner workings, such as the gossip protocol, hinted handoffs, read repairs, Merkle trees, and more.

Chapter 7, Configuring Cassandra

This chapter shows you how to specify partitioners, replica placement strategies, and snitches. We set up a cluster and see the implications of different configuration choices.

Chapter 8, Clients

There are a variety of clients available for different languages, including Java, Python, node.js, Ruby, C#, and PHP, in order to abstract Cassandra's lower-level API. We help you understand common driver features.

Chapter 9, Reading and Writing Data

We build on the previous chapters to learn how Cassandra works "under the covers" to read and write data. We'll also discuss concepts such as batches, lightweight transactions, and paging.

Chapter 10, Monitoring

Once your cluster is up and running, you'll want to monitor its usage, memory patterns, and thread patterns, and understand its general activity. Cassandra has a rich Java Management Extensions (JMX) interface baked in, which we put to use to monitor all of these and more.

Chapter 11, Maintenance

The ongoing maintenance of a Cassandra cluster is made somewhat easier by some tools that ship with the server. We see how to decommission a node, load balance the cluster, get statistics, and perform other routine operational tasks.

Chapter 12, Performance Tuning

One of Cassandra's most notable features is its speed—it's very fast. But there are a number of things, including memory settings, data storage, hardware choices, caching, and buffer sizes, that you can tune to squeeze out even more performance.

Chapter 13, Security

NoSQL technologies are often slighted as being weak on security. Thankfully, Cassandra provides authentication, authorization, and encryption features, which we'll learn how to configure in this chapter.

Chapter 14, Deploying and Integrating

We close the book with a discussion of considerations for planning cluster deployments, including cloud deployments using providers such as Amazon, Microsoft, and Google. We also introduce several technologies that are frequently paired with Cassandra to extend its capabilities.



Cassandra Versions Used in This Book

This book was developed using the Cassandra 3.X series of releases, along with the DataStax Java Driver version 3.0.

When discussing features added in releases 2.0 and later, we cite the release in which the feature was added for readers who may be using earlier versions and are considering whether to upgrade.

New for the Second Edition

The first edition of *Cassandra: The Definitive Guide* was the first book published on Cassandra, and has remained highly regarded over the years. However, the Cassandra landscape has changed significantly since 2010, both in terms of the technology itself and the community that develops and supports that technology. Here's a summary of the key updates we've made to bring the book up to date:

A sense of history

The first edition was written against the 0.7 release in 2010. As of 2016, we're up to the 3.X series. The most significant change has been the introduction of CQL and deprecation of the old Thrift API. Other new architectural features include secondary indexes, materialized views, and lightweight transactions. We provide a summary release history in Chapter 2 to help guide you through the changes.

As we introduce new features throughout the text, we frequently cite the releases in which these features were added.

Giving developers a leg up

Development and testing with Cassandra has changed a lot over the years, with the introduction of the CQL shell (`cqlsh`) and the gradual replacement of community-developed clients with the drivers provided by DataStax. We give in-depth treatment to `cqlsh` in Chapters 3 and 4, and the drivers in Chapters 8 and 9. We also provide an expanded description of Cassandra's read path and write path in Chapter 9 to enhance your understanding of the internals and help you understand the impact of decisions.

Maturing Cassandra operations

As more and more individuals and organizations have deployed Cassandra in production environments, the knowledge base of production challenges and best practices to meet those challenges has increased. We've added entirely new chapters on security (Chapter 13) and deployment and integration (Chapter 14), and greatly expanded the monitoring, maintenance, and performance tuning chapters (Chapters 10 through 12) in order to relate this collected wisdom.

Conventions Used in This Book

The following typographical conventions are used in this book:

Italic

Indicates new terms, URLs, email addresses, filenames, and file extensions.

Constant width

Used for program listings, as well as within paragraphs to refer to program elements such as variable or function names, databases, data types, environment variables, statements, and keywords.

Constant width bold

Shows commands or other text that should be typed literally by the user.

Constant width italic

Shows text that should be replaced with user-supplied values or by values determined by context.



This element signifies a tip or suggestion.



This element signifies a general note.



This element indicates a warning or caution.

Using Code Examples

The code examples found in this book are available for download at <https://github.com/jeffreyscarpenter/cassandra-guide>.

This book is here to help you get your job done. In general, you may use the code in this book in your programs and documentation. You do not need to contact us for permission unless you're reproducing a significant portion of the code. For example, writing a program that uses several chunks of code from this book does not require permission. Selling or distributing a CD-ROM of examples from O'Reilly books does require permission. Answering a question by citing this book and quoting example code does not require permission. Incorporating a significant amount of example code from this book into your product's documentation does require permission.

We appreciate, but do not require, attribution. An attribution usually includes the title, author, publisher, and ISBN. For example: "*Cassandra: The Definitive Guide, Second Edition*, by Jeff Carpenter. Copyright 2016 Jeff Carpenter, 978-1-491-93366-4."

If you feel your use of code examples falls outside fair use or the permission given here, feel free to contact us at permissions@oreilly.com.

Safari® Books Online



Safari Books Online is an on-demand digital library that delivers expert content in both book and video form from the world's leading authors in technology and business.

Technology professionals, software developers, web designers, and business and creative professionals use Safari Books Online as their primary resource for research, problem solving, learning, and certification training.

Safari Books Online offers a range of plans and pricing for enterprise, government, education, and individuals.

Members have access to thousands of books, training videos, and prepublication manuscripts in one fully searchable database from publishers like O'Reilly Media, Prentice Hall Professional, Addison-Wesley Professional, Microsoft Press, Sams, Que, Peachpit Press, Focal Press, Cisco Press, John Wiley & Sons, Syngress, Morgan Kaufmann, IBM Redbooks, Packt, Adobe Press, FT Press, Apress, Manning, New Riders, McGraw-Hill, Jones & Bartlett, Course Technology, and hundreds more. For more information about Safari Books Online, please visit us online.

How to Contact Us

Please address comments and questions concerning this book to the publisher:

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472
800-998-9938 (in the United States or Canada)
707-829-0515 (international or local)
707-829-0104 (fax)

We have a web page for this book, where we list errata, examples, and any additional information. You can access this page at <http://bit.ly/cassandra2e>.

To comment or ask technical questions about this book, send email to bookquestions@oreilly.com.

For more information about our books, courses, conferences, and news, see our website at <http://www.oreilly.com>.

Find us on Facebook: <http://facebook.com/oreilly>

Follow us on Twitter: <http://twitter.com/oreillymedia>

Watch us on YouTube: <http://www.youtube.com/oreillymedia>

Acknowledgments

There are many wonderful people to whom we are grateful for helping bring this book to life.

Thank you to our technical reviewers: Stu Hood, Robert Schneider, and Gary Dusba-bek contributed thoughtful reviews to the first edition, while Andrew Baker, Ewan Elliot, Kirk Damron, Corey Cole, Jeff Jirsa, and Patrick McFadin reviewed the second edition. Chris Judson's feedback was key to the maturation of Chapter 14.

Thank you to Jonathan Ellis and Patrick McFadin for writing forewords for the first and second editions, respectively. Thanks also to Patrick for his contributions to the Spark integration section in Chapter 14.

Thanks to our editors, Mike Loukides and Marie Beaugureau, for their constant support and making this a better book.

Jeff would like to thank Eben for entrusting him with the opportunity to update such a well-regarded, foundational text, and for Eben's encouragement from start to finish.

Finally, we've been inspired by the many terrific developers who have contributed to Cassandra. Hats off for making such an elegant and powerful database.

Table of Contents

Foreword.....	xiii
Foreword.....	xv
Preface.....	xvii
1. Beyond Relational Databases.....	1
What's Wrong with Relational Databases?	1
A Quick Review of Relational Databases	5
RDBMSs: The Awesome and the Not-So-Much	5
Web Scale	12
The Rise of NoSQL	13
Summary	15
2. Introducing Cassandra.....	17
The Cassandra Elevator Pitch	17
Cassandra in 50 Words or Less	17
Distributed and Decentralized	18
Elastic Scalability	19
High Availability and Fault Tolerance	19
Tuneable Consistency	20
Brewer's CAP Theorem	23
Row-Oriented	27
High Performance	28
Where Did Cassandra Come From?	28
Release History	30
Is Cassandra a Good Fit for My Project?	35
Large Deployments	35