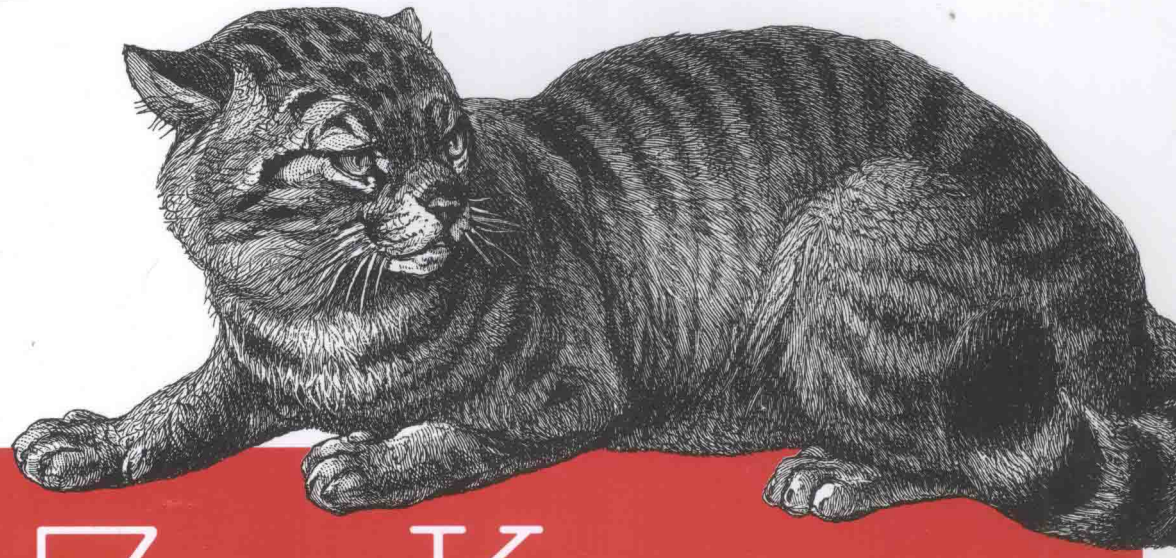
O'REILLY®

HZ BOOKS
华章IT



ZooKeeper

分布式过程协同技术详解

ZooKeeper

Flavio Junqueira Benjamin Reed 著
谢超 周贵卿 译

机械工业出版社
China Machine Press

ZooKeeper
分布式过程协同技术详解

Flavio Junqueira, Benjamin Reed 著
谢超 周贵卿 译

Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo O'REILLY®
O'Reilly Media, Inc.授权机械工业出版社出版

机械工业出版社

图书在版编目（CIP）数据

ZooKeeper：分布式过程协同技术详解/［美］荣凯拉（Junqueira, F.），［美］里德（Reed, B.）著；谢超，周贵卿 译. —北京：机械工业出版社，2016.1
（O'Reilly精品图书系列）
书名原文：ZooKeeper: Distributed Process Coordination
ISBN 978-7-111-52431-1

I. Z… II. ①荣… ②里… ③谢… ④周… III. ①分布式操作系统－研究 IV. TP316.4
中国版本图书馆CIP数据核字（2015）第309542号

北京市版权局著作权合同登记
图字：01-2014-2012号

书　　名/ ZooKeeper：分布式过程协同技术详解
书　　号/ ISBN 978-7-111-52431-1
责任编辑/ 秦健
封面设计/ Randy Comer，张健
出版发行/ 机械工业出版社
地　　址/ 北京市西城区百万庄大街22号（邮政编码 100037）
印　　刷/ 北京市荣盛彩色印刷有限公司
开　　本/ 178毫米×233毫米　16开本　14印张
版　　次/ 2016年1月第1版　2016年1月第1次印刷
定　　价/ 69.00元（册）

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
读者信箱：hzit@hzbook.com
购书热线：(010)68326294；88379649；68995259
投稿热线：(010)88379604
客服热线：(010)88379426；88361066

O'Reilly Media, Inc.介绍

O'Reilly Media通过图书、杂志、在线服务、调查研究和会议等方式传播创新知识。自1978年开始，O'Reilly一直都是前沿发展的见证者和推动者。超级极客们正在开创着未来，而我们关注真正重要的技术趋势——通过放大那些“细微的信号”来刺激社会对新科技的应用。作为技术社区中活跃的参与者，O'Reilly的发展充满了对创新的倡导、创造和发扬光大。

O'Reilly为软件开发人员带来革命性的“动物书”；创建第一个商业网站（GNN）；组织了影响深远的开放源代码峰会，以至于开源软件运动以此命名；创立了Make杂志，从而成为DIY革命的主要先锋；公司一如既往地通过多种形式缔结信息与人的纽带。O'Reilly的会议和峰会集聚了众多超级极客和高瞻远瞩的商业领袖，共同描绘出开创新产业的革命性思想。作为技术人士获取信息的选择，O'Reilly现在还将先锋专家的知识传递给普通的计算机用户。无论是通过书籍出版，在线服务或者面授课程，每一项O'Reilly的产品都反映了公司不可动摇的理念——信息是激发创新的力量。

业界评论

“O'Reilly Radar博客有口皆碑。”

——Wired

“O'Reilly凭借一系列（真希望当初我也想到了）非凡想法建立了数百万美元的业务。”

——Business 2.0

“O'Reilly Conference是聚集关键思想领袖的绝对典范。”

——CRN

“一本O'Reilly的书就代表一个有用、有前途、需要学习的主题。”

——Irish Times

“Tim是位特立独行的商人，他不光放眼于最长远、最广阔的视野并且切实地按照Yogi Berra的建议去做了：‘如果你在路上遇到岔路口，走小路（岔路）。’回顾过去Tim似乎每一次都选择了小路，而且有几次都是一闪即逝的机会，尽管大路也不错。”

——Linux Journal

译者序

摩尔定律揭示了集成电路每18个月计算性能就会增加一倍。随着信息的飞速膨胀，很多应用都无法依赖单个服务器的性能升级来处理如此庞大的数据量，分布式系统和应用越来越受到人们的青睐。分布式系统和应用不仅能提供更强的计算能力，还能为我们提供更好的容灾性和扩展性。

在实际开发分布式应用时，开发人员与运维人员都会花费大量时间和精力来处理异构系统中的协作通信问题。这也许并不是你想要的，你最关心的是战略业务是否正常，能否更快更好地提供自己主营业务的系统和服务。因此对分布式系统的协作处理上，需要专门处理协作问题的系统来帮助我们。

ZooKeeper是Google的Chubby项目的开源实现，它曾经作为Hadoop的子项目，在大数据领域得到广泛应用。ZooKeeper以Fast Paxos算法为基础，同时为了解决活锁问题，对Fast Paxos算法进行了优化，因此也可以广泛用于大数据之外的其他分布式系统，为大型分布式系统提供可靠的协作处理功能。比如小米公司的米聊，其后台就采用了ZooKeeper作为分布式服务的统一协作系统。而阿里公司的开发人员也广泛使用ZooKeeper，并对其进行了适当修改，开源了一款TaoKeeper软件，以适应自身业务需要。

本书首先从分布式系统的基本概念入手，然后介绍实际开发编程的接口和技巧，最后谈及运维人员所关心的配置维护知识。翻译过程中，译者对原版书籍通读一遍，对ZooKeeper又有了新的认识和理解，获得了分布式应用构建中需要注意的很多细节，这本书可谓是实际开发和维护中的一本最佳参考书籍。对于这么优秀的一本书，翻译时译

者惶恐于译文对读者理解的影响，尽最大努力保持原文意思，以便读者真正能够领悟ZooKeeper的精髓。

由于译者水平有限，译文中的不当之处在所难免，恳请广大读者批评指正。

谢超　周贵卿

目录

第三部分 ZooKeeper的管理

前言

构建分布式系统并不容易。然而，人们日常所使用的应用大多基于分布式系统，在短时间内依赖于分布式系统的现状并不会改变。Apache ZooKeeper旨在减轻构建健壮的分布式系统的任务。ZooKeeper基于分布式计算的核心概念而设计，主要目的是给开发人员提供一套容易理解和开发的接口，从而简化分布式系统构建的任务。

即使有了ZooKeeper，但开发中分布式处理的环节并不是微不足道的事情，因此我们编写了这本书，通过这本书可以让你快速熟悉如何通过Apache ZooKeeper构建分布式系统。我们从基本的概念入手，这样可以使你觉得自己就像是分布式系统的专家一样，在你看到一系列需要注意的警告时，你可能会有一些沮丧，不过不用担心，如果你能够很好地理解我们所阐述的关键点，你已经走在构建良好的分布式系统的正确道路上了。

目标读者

本书适用于分布式系统的开发人员，以及使用ZooKeeper进行生产经营的应用程序运维人员。我们假设读者具备Java语言的知识，并且本书为读者提供了关于分布式系统中概念的大量背景知识，以便你更好地使用ZooKeeper。

本书内容介绍

第一部分阐述了Apache ZooKeeper这类系统的设计目的和动机，并介绍分布式系统的一些必要背景知识。

- 第1章介绍了ZooKeeper可以做什么，以及其设计如何支撑这些任务。
- 第2章介绍了基本概念和基本组成模块，并通过命令行工具的具体操作介绍ZooKeeper可以做什么。

第二部分阐述程序员所需要掌握的ZooKeeper库调用方法和编程技巧，虽然对系统运维人员来说也有一定价值，但也可以不选择阅读。这一部分主要以Java语言的API为主，因为Java是非常流行的开发语言，如果你之前使用其他开发语言，可以通过这一部分内容来学习基本的技术和方法调用，之后通过其他语言来实现。另外，我们也为C语言的应用开发人员提供了一章内容的开发方法。

- 第3章介绍Java语言的API。
- 第4章解释如何跟踪和处理ZooKeeper中的状态变更情况。
- 第5章介绍如何在系统或网络故障时恢复应用。
- 第6章介绍为了避免故障要注意的一些繁杂却很重要的场景。
- 第7章介绍C语言版的API，该章也可以作为非Java语言实现的ZooKeeper API的基础，对非Java语言的开发人员非常有帮助。
- 第8章介绍一款更高层级的封装的ZooKeeper接口。

第三部分主要适用于ZooKeeper的系统运维人员，尤其在第9章中，即便对开发人员也很有价值。

- 第9章介绍ZooKeeper的作者们在设计时所采用的方案，这些知识对运维管理非常有帮助。
- 第10章介绍如何对ZooKeeper进行配置。

本书约定

本书中采用了以下排版约定：

斜体

用于重点介绍新的术语、URL、命令、工具组件以及文件和目录名称。

等宽字体（`Constant width`）

指示变量、方法、类型、参数、对象以及其他代码结构。

等宽加粗（**`Constant width bold`**）

指示需要用户输入的命令或其他文本信息，同时也用于命令输出中的重要信息。

等宽斜体（*Constant width italic*）

指示代码或命令中的占位符，这些占位符需要在实际中替换为合适的值。

注意：表示一些小窍门、建议或普通注解。

示例代码

代码、练习等附加资料可以到O'Reilly官方网站本书页面下载。

本书用于协助读者构建系统。一般而言，如果本书提供了示例代码，你可以在自己的程序或文档中使用，并不需要联系我们获得授权，除非你复制了大量代码。例如，你在开发程序时使用了本书中的好几处代码则不需要授权，若以CD-ROM方式出售并发布O'Reilly书籍中的示例则需要得到授权许可，引用本书及其示例代码来解答问题并不需要授权许可，将本书中大量示例代码引入你自己的著作中则需要授权许可。

我们非常感谢各类引文参考，但并不强制约束。引文参考包括书名、作者、出版方和ISBN。例如："*ZooKeeper* by Flavio Junqueira and Benjamin Reed (O'Reilly). Copyright 2014 Flavio Junqueira and Benjamin Reed, 978-1-449-36130-3."

如果你对合理使用示例代码时有疑问，或对以上所介绍的许可授权有疑问，请通过*permissions@oreilly.com*联系我们。

联系我们

有关本书的任何建议和疑问，可以通过下列方式与我们取得联系：

美国：

O'Reilly Media，Inc.
1005 Gravenstein Highway North
Sebastopol，CA 95472

中国：

北京市西城区西直门南大街2号成铭大厦C座807室（100035）
奥莱利技术咨询（北京）有限公司

我们会在本书的网页中列出勘误表、示例和其他信息。可以通过访问如下网址获得：

http://shop.oreilly.com/product/0636920028901.do

要评论或询问本书的技术问题，请发送电子邮件到：

bookquestions@oreilly.com

想了解关于O'Reilly图书、课程、会议和新闻的更多信息，请访问以下网站：

http://www.oreilly.com.cn
http://www.oreilly.com

致谢

我们要感谢本书的编辑人员，从最初的Nathan Jepson到后来Andy Oram的努力，他们的出色工作让我们得以出版此书。

我们要感谢所有在此书上花费这么多时间的家人和雇主，希望你也能欣赏我们的成果。

我们要感谢为本书花费大量时间的审阅者们，他们给予我们很大帮助，为我们提供建议来改进本书，他们包括Patrick Hunt、Jordan Zimmerman、Donald Miner、Henry Robinson、Isabel Drost-Fromm和Thawan Kooburat。

ZooKeeper是Apache ZooKeeper社区共同创作的，我们与这些杰出的提交者和贡献者一同工作，非常荣幸能与他们一同工作。我们还要向ZooKeeper的用户们致谢，多年以来，他们向我们提交bug，给予我们很多反馈信息和鼓励。

第一部分

ZooKeeper的概念和基础

这一部分适合任何对ZooKeeper感兴趣的读者，该部分介绍ZooKeeper所处理的问题，以及在ZooKeeper的设计中的权衡取舍。