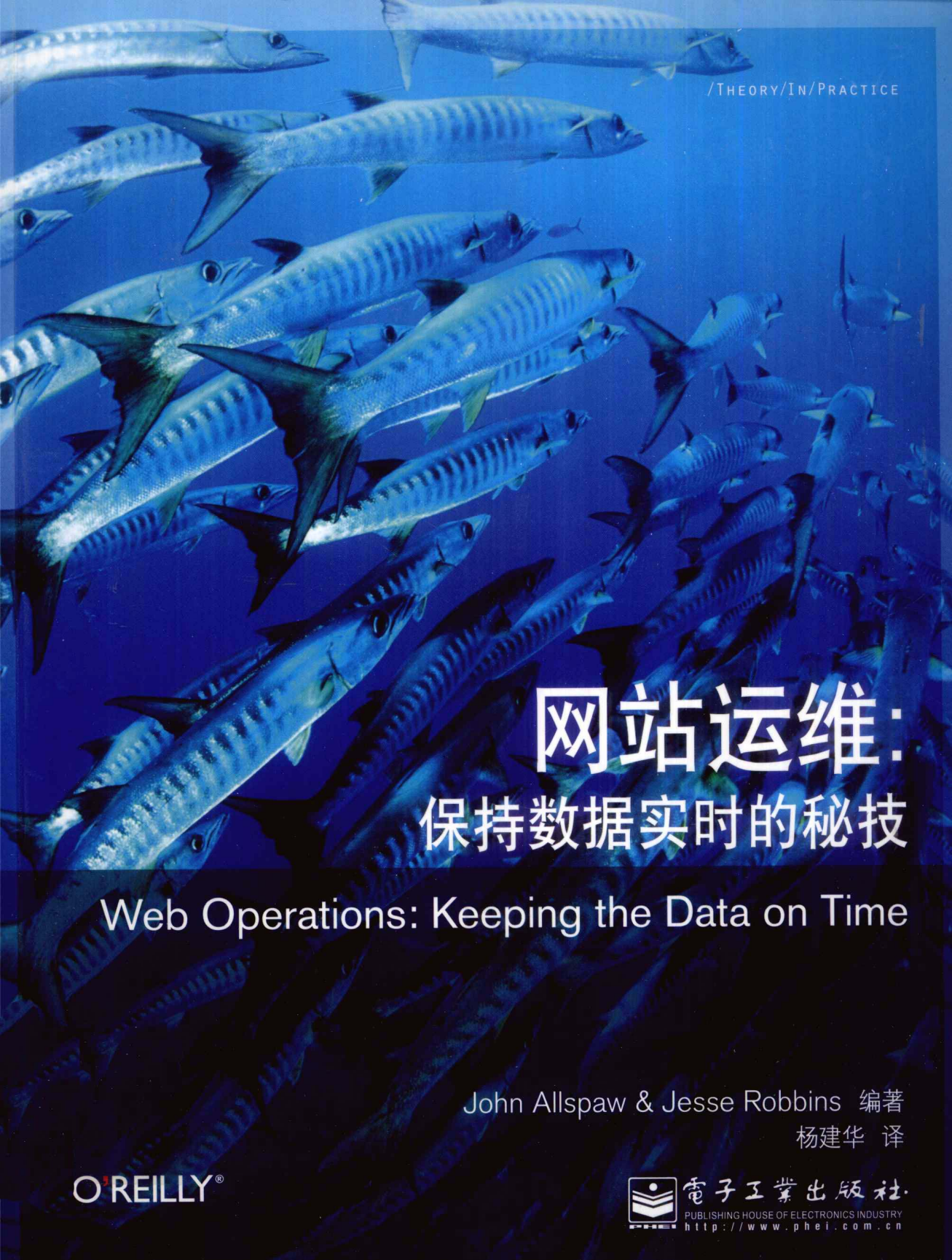




/THEORY/IN/PRACTICE

网站运维： 保持数据实时的秘技

Web Operations: Keeping the Data on Time

John Allspaw & Jesse Robbins 编著
杨建华 译

O'REILLY®



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

O'REILLY®

网站运维：

保持数据实时的秘技

Web Operations: Keeping the Data on Time

John Allspaw & Jesse Robbins 编著
杨建华 译

电子工业出版社
Publishing House of Electronics Industry
北京•BEIJING

内 容 简 介

本书的各位合作者以自己的亲身经历,从不同的侧面讲述了一个正在发展的新兴技术领域——网站运维:职业特点、技术架构、测量与监控、开发与部署、用户体验、后端存储,以及如何应对流量激增、如何优雅地失败,特别是如何将软件开发领域中的敏捷方法应用于网站运维。阅读本书,读者不仅能够学到一些具体的技术,还能够开阔眼界,因为书中不仅有具体方法的讨论,更有基于十多年运维经验的哲理思考。

本书适合网站架构师、开发人员、运维人员(包括但不限于数据库管理员、系统管理员、网络管理员、发布管理员等)阅读,同时也适合项目经理等管理人员阅读。

978-1-449-37744-1 Web Operations: Keeping the Data on Time ©2010 by O'Reilly Media, Inc. Simplified Chinese edition, jointly published by O'Reilly Media, Inc. and Publishing House of Electronics Industry, 2011. Authorized translation of the English edition, 2010 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.

本书中文简体版专有版权由O'Reilly Media, Inc.授予电子工业出版社,未经许可,不得以任何方式复制或抄袭本书的任何部分。

版权贸易合同登记号 图字:01-2011-4340

图书在版编目(CIP)数据

网站运维:保持数据实时的秘技/(美)阿尔斯帕瓦(Allspaw,J.), (美)罗宾斯(Robbins,J.)编著;杨建华译. —北京:电子工业出版社, 2011.9

书名原文: Web Operations: Keeping the Data on Time

ISBN 978-7-121-14163-8

I. ①网… II. ①阿…②罗…③杨… III. ①网站—开发 IV. ①TP393.092

中国版本图书馆CIP数据核字(2011)第146580号

策划编辑:刘 皎

责任编辑:贾 莉

封面设计: Karen Montgomery

印 刷:

装 订: 北京中新伟业印刷有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路173信箱 邮编: 100036

开 本: 787×980 1/16

印张: 19 字数: 380千字

印 次: 2011年9月第1次印刷

定 价: 49.00元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至zlt@phei.com.cn,盗版侵权举报请发邮件至dbqq@phei.com.cn。

服务热线:(010)88258888。

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

序

世界上第一批网站真正达到规模以来，已经过去十多年了。我们就是当年的运维工程师，在那些日子里，见证了网站前所未有的快速增长，过去也没有人知道该怎样管理这样的网站。而保持系统正常运行正是我们的责任。

当其他人在发布会晚宴上庆祝时，我们却在数据中心拆装最后的服务器。忙完以后，已是深夜了，在桌边坐下来，看着日志文件和访问曲线不断增长，我们个个都容光焕发。

我们的经历是有普遍意义的：软件崩溃了，或不能随着规模的增长而缩放。数据库崩溃了，数据也被污染了，而服务器、磁盘、交换机，这些产品的制造商都保证说绝对不会出问题，却恰恰都出问题了。网站遭到了黑客攻击——开始时是为了好玩，后来就为钱而干了。我们刚把事情搞定，这时又要推出一项新功能，导致访问量急剧上升，所有努力又全完了。

在那些早期的日子里，因为没有预算，我们只能使用我们能找到的东西。我们从使用那些杂物间里找出来的机器，逐渐增长到分布在全球的兆瓦规模的数据中心，这些数据中心里的机器仍然是我们能找到的最便宜的机器。

随着网站逐渐走向规模化，现实世界及其危险也就逐渐凸显了。数据中心会遭遇火灾、洪水、被飓风吹得一片狼藉。电力可能中断，备用发电机可能不起作用——或可以启动，但马上又没油了——或被什么人揪下了紧急断电（Emergency Power Off）按钮。制冷系统会损坏，洒水装置会漏水，光缆可能被挖土机、松鼠或在海底爬行的不知什么生物给切断。

人、机器、大自然会在我们能够想象到的所有方面对我们发起挑战，然后在我们永远也想不到的方面使我们目瞪口呆。

呼机鸣叫、朋友无心地询问“网站宕了吗？”或CEO惊恐与愤怒地叫喊时，我们立马投入工作。我们总是第一个知道网站宕机的，网站正常后，也总是最后一个离开的。

总是这样。

每天，我们都能更机敏一点，更聪明一点，也总能学到一些小技巧。十多年前我们写的脚本，现在已经发展为成熟的工具和语言了，而且围绕这些已经出现了完整的行业，知识、经验、工具、过程正在成长成为一种技艺，我们称其为“Web运维”。

我们之所以称Web运维是一种技艺，而不是科学，是有原因的。在这方面没有标准、认证、甚至正式的学校教育（至少目前是这样）。我们花了很长时间才学到了怎么做，又花了更长时间才算精通。每个人在每个阶段都必须找到他或她自己的风格。这里没有“正确答案，”有的只是什么管用（现在），以及下次要做得更好的承诺。

Web正在改变我们的生活方式，并且触及到了每一个人。随着越来越多的人依赖于Web，他们最终将依赖于我们。

Web运维就是这样的工作。

——Jesse Robbins

本书的合作者已经将他们的收入捐献给了826基金会，以帮助孩子们在如Superhero Supply Company、Greenwood Space Travel Supply Company以及Liberty Street Robot Supply & Repair Shop这样的地方学习对读书的热爱。

前言

设计、构建、维护一个不断增长的网站，在其进入系统管理和软件开发阶段后，将有很多不同寻常的挑战。其中之一，就是Web从来不睡觉。网站是全球可访问的，所以不存在“适宜”的时间来做改变、更新、维护，只有一些“不适宜”的时间，这也意味着中断肯定影响某些在线的人，而不管是什么时间。

随着Web应用逐渐进入我们的日常生活，Web应用也就变得越来越复杂。随着复杂性而来的是，更多的部件要构建与维护，不幸的是，同时也就会有更多部件出问题。除此之外，还要保证访问更快速、更安全，而且全球通。所有这些事情加起来，就构成了一个特殊的工程领域：Web运维。

Jesse Robbins和我构想了这样一本书，目的是将一些业内专家的见识整理出来，为这个仍然在发展的领域提供一些参考。我们列出的题目仅是冰山一角，这些专家将自己的呕心沥血化为真知灼见，将当年的战场鏖战化为轻松故事，一一道来，以飨读者。

本书的组织

本书的章节安排如下：

第1章，作为职业的Web运维，Theo Schlossnagle，描述了这个领域都包含些什么，并且强调了所需技能如何由经验所得，而非正规教育。

第2章，Picnik如何应用云计算：所学到的教训，Justin Huff，解释了Picnik.com如何部署并维持一个建立在自有硬件（on-premises hardware）和云服务（cloud services）混合上的基础架构的。

第3章，基础架构与应用程序测量，Matt Massie和我，讨论了从应用程序和基础架构获取测量数据的重要性，以及如何获取这些数据。

第4章，连续部署，Eric Ries，对连续、小批量部署代码的优点，给出了他的看法。

- 第5章，作为代码的基础架构，Adam Jacob，简要介绍了关于配置与部署管理的理论与方法。
- 第6章，监控，Patrick Debois，讨论了设计一个监控系统时需要考虑的各种问题。
- 第7章，复杂系统是如何失败的，是Richard Cook博士的一篇白皮书，讨论了系统失败及复杂性的本质，这种复杂性经常出现在Web结构上。这里，Richard Cook博士还增加了一些与Web运维有关的注释。
- 第8章，社区管理与Web运维，是我和Heather Champ的一篇访谈，话题是如何在人这方面处理停机与降级运行的问题。
- 第9章，处理非预期的访问量激增，Brian Moon，讨论了Dealnews.com处理访问量激增问题的经验。
- 第10章，开发者与运维者的协调与合作，Paul Hammond，列出了一些不论从技术上还是文化上，都需要开发者与运维者合作的地方。
- 第11章，你的访问者感觉怎么样：面向用户的测量，Alistair Croll和Sean Power，讨论了如何将测量用于度量网站的用户体验。
- 第12章，将关系数据库用于Web的战略战术，Baron Schwartz，展示了架构数据库的常用方法，以及随着规模增长会遇到的一些问题。
- 第13章，如何优雅地失败：事后处理的艺术与科学，Jake Loomis，讨论了如何做成（或做不成）一个好的事后处理及根本原因分析过程。
- 第14章，存储，Anoop Nagwani，一个不断增长的Web应用的存储系统应该如何设计与维护？本章讨论了相关方法与需要考虑的问题。
- 第15章，非关系数据库，Eric Florenzano，“非关系”数据库技术正在大量涌现，该部分列举了该类技术的一些优点及需要考虑的问题。
- 第16章，敏捷基础架构，Andrew Clay Shafer，讨论了运维中人和过程的因素，以及如何将敏捷的理念和方法应用（或不用）到运维中。
- 第17章，夜半鬼怪（以及如何依然安睡），Mike Christian，讨论了不同级别的可用性 & 业务持续规划（Business Continuity Planning, BCP）的方法及危险。

本书对象

本书的读者对象是开发者、系统管理员，以及数据库、网络或其他所有运维Web应用的工程师。本书所讨论的内容都适用于Web运维这个仍然在继续发展的领域。

本书所用的约定

本书使用下列印刷约定：

斜体 (Italic)

表示URL。

等宽字体 (Constant width)

用于程序列表，以及在段落内表示程序元素（如变量或函数名、数据库、数据类型、环境变量、语句及关键字等）。

使用代码样例

本书是帮助你工作的，一般情况下，你可以在程序和文档中使用本书中的代码，而不需要得到我们的允许，除非你要把大量的代码拿去再生产。譬如，用本书中的一些代码写一个程序，不需要得到允许，但出售或分发O'Reilly代码样例的CD-ROM则需要得到允许。由于回答一个问题而引用本书或书中的代码样例，不需要得到允许。将本书中大量代码样例编入你的产品文档，则需要得到允许。

我们欣赏写明归属，但并不要求。归属通常包括书名、作者、出版商及ISBN，如“*Web Operations: Keeping Data On Time*, edited by John Allspaw and Jesse Robbins. Copyright 2010 O'Reilly Media, Inc., 978-1-449-37744-1.”等。

要是你觉得你对代码样例的使用超出了合理使用或上面所允许的范围，请通过 permissions@oreilly.com 联系我们。

如何联系我们

请将对本书的评价和存在的问题通过如下地址告知出版者：

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

O'Reilly的每一本书都有专属网站，你可以在那找到关于本书的相关信息，包括勘误列表、示例代码以及其他的信息。本书的网站地址是：

<http://www.oreilly.com/catalog/9781449377441>

对于本书的评论和技术性的问题，请发送电子邮件到：

bookquestions@oreilly.com

关于本书的更多信息、会议、资料中心和O'Reilly网络，请访问以下网站：

<http://www.oreilly.com/>

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

Safari图书在线

Safari[®]

Safari图书在线是一个按需服务的数字图书馆，可以很容易地搜索超过7500种的技术、参考书及视频，从中快速获取答案。

只要有一个订阅，你就可以在我们的图书馆中在线阅读任何书页、观看任何视频。可以在手机或移动设备上读书，在书籍完成送去印刷之前获得新书目录，在书籍开发阶段获取对手稿的排他性访问权并给作者提供反馈，复制与粘贴代码样例、组织你的收藏夹、下载章节、对关键段落做书签标记、创建注释、打印书页，以及其他各种好处。

O'Reilly Media已经将本书英文版上传到Safari图书在线，要访问这本书或O'Reilly及其他出版商的类似主题，请免费注册：<http://my.safaribooksonline.com>。

致谢

John Allspaw向Elizabeth、Sadie和Jack表达谢意，感谢你们在我忙于这本书时表现出的极大耐心。我还要感谢所有的作者，当然白天他们都有自己的工作，但在稿期很紧的情况下，却都能够按时交稿。

Jesse Robbins向John Allspaw表示感谢，大部分工作都是他做的，没有他，永远也不会有这本书。

目录

序	xi
前言	xiii
第1章 作为职业的Web运维	1
<i>Theo Schlossnagle</i>	
为什么Web运维如此艰难?	1
从学徒到师傅	4
结语	9
第2章 Picnik如何应用云计算：所学到的教训	10
<i>Justin Huff</i>	
什么地方适合云计算（以及为什么！）	11
什么地方不适合云计算（对Picnik而言）	17
结语	18
第3章 基础架构与应用程序测量	19
<i>John Allspaw, Matt Massie</i>	
时间分辨率和存留时间的考虑	20
测量数据采集与存储的地点	21
测量数据的层次	22
为异常检测和报警提供环境	25
日志记录也是测量数据	26
将变化管理和事件的时间线建立关联	27
给测量数据加入报警机制	28
使用测量数据建立加载-反馈机制	29

展示一个测量数据采集系统：Ganglia	32
结语	43
第4章 连续部署	44
<i>Eric Ries</i>	
小批量意味着更快的反馈	44
小批量意味着问题即刻被本地化	44
小批量能够减少风险	45
小批量可以降低总开销	45
质量卫士的挽歌	47
让我们开始吧	50
连续部署用于关键任务应用	54
结语	57
第5章 作为代码的基础架构	58
<i>Adam Jacob</i>	
面向服务体系结构	60
结语	71
第6章 监控	72
<i>Patrick Debois</i>	
故事：“旅程的开端”	72
步骤1：理解你在监控什么	76
步骤2：理解正常行为	84
步骤3：有备而学	90
结语	93
第7章 复杂系统是如何失败的	94
<i>John Allspaw和Richard Cook</i>	
复杂系统是如何失效的	94
进一步的读物	101

第8章 社区管理与Web运维	103
<i>Heather Champ和John Allspaw</i>	
第9章 处理非预期的访问量激增	112
<i>Brian Moon</i>	
一切是如何开始的	112
警报连连	113
扑灭烈火	114
周末逃生	115
未雨绸缪	116
救命稻草CDN	116
代理服务器	116
围剿踩踏	117
将代码基流水化	118
我们怎么知道它能否工作?	119
真实测试	120
学到的教训	120
自那以来的改进	121
第10章 开发者与运维者的协调与合作	122
<i>Paul Hammond</i>	
部署	123
共享、开放的基础架构	126
信任	128
随叫随到的开发人员	131
避免指责	135
结语	137
第11章 你的访问者感觉怎么样：面向用户的测量	139
<i>Alistair Croll和Sean Power</i>	
为什么要采集面向用户的测量数据?	140

是什么使网站变得很慢?	144
测量延迟	147
编写SLA	153
访客结果: 分析	155
市场营销关心的其他测量数据	160
用户体验如何影响Web运维	161
Web监控的未来	162
结语	167
第12章 将关系数据库用于Web的战略战术	169
<i>Baron Schwartz</i>	
Web数据库需求	170
典型的Web数据库是如何增长的	175
对集群的渴望	181
数据库战略	186
数据库战术	193
结语	198
第13章 如何优雅地失败: 事后处理的艺术与科学	200
<i>Jake Loomis</i>	
最糟的事后分析	200
什么是事后分析?	201
什么时候引入事后分析	203
邀请谁参加事后分析	204
进行事后分析	204
事后分析的后续工作	205
结语	207
第14章 存储	208
<i>Anoop Nagwani</i>	
数据资产的库存	208

数据保护	211
容量规划	218
存储大小的变化	219
运维	221
结语	223
第15章 非关系数据库	224
<i>Eric Florenzano</i>	
NoSQL数据库概览	225
某些系统细节	228
结语	238
第16章 敏捷基础架构	239
<i>Andrew Clay Shafer</i>	
敏捷基础架构	241
那么，问题是什么？	244
兴趣与实践的社区	253
贸易区和道歉	253
结语	256
第17章 夜间鬼魅（以及如何高枕无忧）	257
<i>Mike Christina</i>	
术语	258
多少个9？	259
影响持续时间对事件持续时间	260
数据中心数量（footprint）	261
逐渐失效	262
不信赖任何人	263
故障转移测试	264
监控和历史模式	264

高枕无忧	265
合作者	267
索引	271