HZ BOOKS
华章科技

世界级日志管理与分析专家亲笔撰写，完美诠释有效的日志分析策略以及最佳实践。

从日志的基本概念到系统运营，从传统的syslog到云计算和大数据环境下的新兴日志分析技术，全面解析日志管理和分析方面的各种实用技术及工具。

ELSEVIER
爱思唯尔

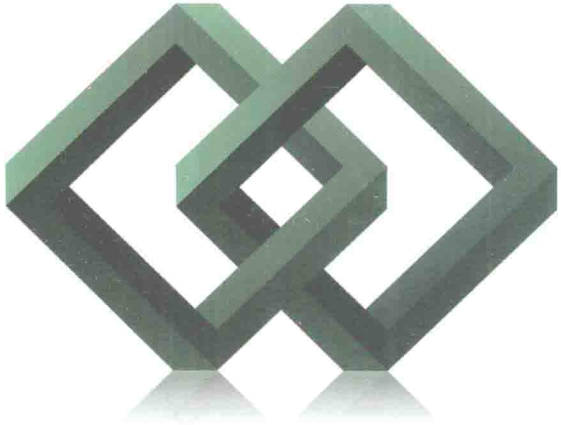
Logging and Log Management

The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management

日志管理与分析权威指南

（美）Anton A. Chuvakin　Kevin J. Schmidt　Christopher Phillips　著

姚军　简于涵　刘晖　等译



机械工业出版社
China Machine Press

华章程序员书库

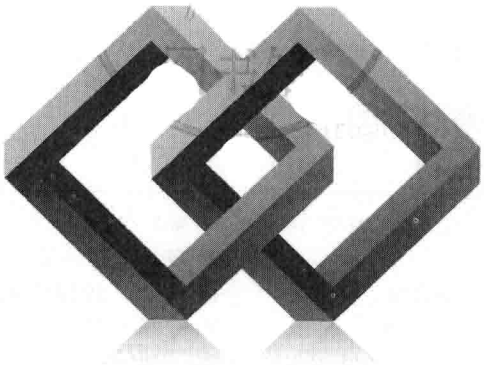
Logging and Log Management

The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management

日志管理与分析权威指南

（美）Anton A. Chuvakin　Kevin J. Schmidt　Christopher Phillips　著

姚军　简于涵　刘晖　等 译



机械工业出版社
China Machine Press

图书在版编目（CIP）数据

日志管理与分析权威指南 /（美）褚瓦金（Chuvakin, A. A.）等著；姚军等译 . —北京：机械工业出版社，2014.6
（华章程序员书库）
书名原文：Logging and Log Management: The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management

ISBN 978-7-111-46918-6

I. 日… II. ①褚… ②姚… III. 计算机网络管理－指南 IV. TP393.07-62

中国版本图书馆 CIP 数据核字（2014）第 116145 号

本书版权登记号：图字：01-2013-6487

日志管理与分析权威指南

[美] Anton A. Chuvakin 等著

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）
责任编辑：吴 怡 责任校对：董纪丽
印 刷：三河市宏图印务有限公司 版 次：2014 年 6 月第 1 版第 1 次印刷
开 本：186mm × 240mm 1/16 印 张：21
书 号：ISBN 978-7-111-46918-6 定 价：69.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991 88361066 投稿热线：（010）88379604
购书热线：（010）68326294 88379649 68995259 读者信箱：hzjsj@hzbook.com

The Translator's Words 译者序

对于广大IT工作者（尤其是负责运维工作的人）来说，日志是一个熟悉的名词，机房中的各种系统、防火墙、交换机、路由器，都在不断地产生日志，而在我们的规章制度上，还有着手工记录或者由某些代理程序采集的运维日志。但是，就是这个天天伴随我们的概念，我们真的熟悉吗？我们是否天天都会去查看重要系统的日志，在系统出现问题或者问题隐患的时候，我们是否能从日志的分析中看出端倪？当系统遭受严重的攻击和破坏时，我们的日志记录系统是否能够幸免于难，帮助我们一步步走出泥淖？

IT安全业界的无数实践告诉我们，健全的日志记录和分析系统是系统正常运营与优化以及安全事故响应的基础，虽然安全系统供应商为我们提供了五花八门的解决方案，但是最终的基础仍是具有充足性、可用性、安全性的日志记录系统。作为运维人员，以及为企业开发应用程序（日志生产者）的开发人员，能不能回答上述的问题，正是对我们实际工作能力的一个考验。

在实际工作中，我们可以看到，许多单位内部对日志并没有充分的认识，安全工作更多地投入设备，比如防火墙、IDS、IPS、防病毒软件等，被动地希望这些系统帮助我们完成一切工作，但是俗话说得好："道高一尺，魔高一丈"，以特征码和预定义规则为基础的上述设备，在防护方面永远落在攻击者的后面，防微杜渐才是真正的出路。作为一名合格的运维人员，了解日志的概念，了解日志的配置和分析方法，是发现威胁、抵御攻击的重要技能，有了这方面的深刻认识，各种自动化安全解决方案才能真正地发挥效能。

多年以来，许多安全书籍都会或多或少地提及日志的作用，但是我们都苦于没有一本全面介绍日志的书籍，这也给人们在日常工作中开发有效的日志记录、分析、响应系统，以及运营中相关的规程带来了困难。

本书是由三位业界资深的安全专家编撰，从日志的基本概念开始，由浅入深地讲述了整个日志生命期的详细过程，书中以大量实例介绍了许多日志方面的最佳实践，从传统的

syslog，一直到正在不断兴起的云计算和大数据，涵盖了这一重要领域中的各项法则，并结合业界中最为常见的监管规章。不仅从详细的技术及工具上，而且从整个运营规程、策略上形成了完整的系统，从而突破了行业和具体软硬件配置的限制，不管读者身处何种规模、何种软硬件配置，均能从本书介绍的概念和思路中获益，并通过自己的努力，形成基于标准、适合自身特点的日志运营架构。

本书涵盖的内容极广，给翻译工作带来了很大的挑战，同时也让译者受益匪浅，我们衷心地希望书中的内容能给广大读者带来帮助，由于译者水平所限，错误在所难免，期待读者朋友们的指正。本书的翻译工作主要由姚军完成，简于涵、刘晖、徐锋、陈绍继、郑端、吴兰陟、施游、林起浪、陈志勇、刘建林、白龙、林耀成、陈霞、方翊、宁懿等人也为本书的翻译工作做出了贡献，在此衷心感谢华章公司的编辑们对翻译工作提出的中肯意见，以及其他工作人员对本书付出的艰辛努力。

About the Author 作者简介

Anton A. Chuvakin 博士是日志管理、SIEM 和 PCI DSS 依从性领域公认的安全专家，他参与撰写了《Security Warrior》(ISBN: 978-0-596-00545-0) 和《Know Your Enemy: Learning About Security Threats》第 2 版（ISBN: 978-0-321-16646-3）、《Information Security Management Handbook》 第 6 版（ISBN: 978-0-8493-7495-1）、《Hacker's Challenge 3:20 Brand-New Forensic Scenarios & Solutions》（ISBN: 978-0-072-26304-6）、《OSSEC Host-Based Intrusion Detection Guide》(Syngress，ISBN: 978-1-59749-240-9）等书籍。

Anton 已经发表了数十篇有关日志管理、关联分析、数据分析、PCI DSS、安全管理等安全主题的文章。他的博客 www.securitywarrior.org 是该领域中最受欢迎的博客之一。此外，Anton 在全球的许多安全会议上发表演讲，包括美国、英国、新加坡、西班牙、俄罗斯等地。他参与新兴的安全标准的制定，并且担任多家安全领域创业公司的顾问。

目前，他运营自己的顾问公司 Security Warrior。在此之前，他曾经是 Qualys 的 PCI 依从性解决方案主管和 LogLogic 的首席日志管理者，任务是为全世界提供关于安全、标准化和运营日志的重要性的培训。在 LogLogic 之前，他曾经受雇于一家安全供应商，担任战略产品管理职务。Anton 拥有 Stony Brook 大学的博士学位。

Kevin J. Schmidt 是 Dell SecureWorks 公司的高级经理，这家业界领先的安全托管服务提供商（MSSP）是 Dell 的下属公司。他负责公司 SIEM 平台主要部分的设计和开发，包括数据获取、关联分析和日志数据分析。就职于 SecureWorks 之前，Kevin 为 Reflex Security 工作，致力于 IPS 引擎和反病毒软件。在此之前，他是 GuradedNet 公司的首席开发人员和架构师，该公司构建了行业最早的 SIEM 平台之一。他还是美国海军预备队（USNR）的军官。Kevin 在软件开发和设计领域有 19 年的经验，其中 11 年从事网络安全领域的研发工作。他持有计算机科学学士学位。

Christopher Phillips 是 Dell SecureWorks 的经理和高级软件开发人员，负责公司 Threat

Intelligence 服务平台的设计和开发。他还负责一个团队，致力于集成来自许多第三方提供商的日志和事件信息，帮助客户通过 Dell SecureWorks 系统和安全专业人士分析信息。在就职于 Dell SecureWorks 之前，他为 McKesson 和 Allscripts 工作，帮助客户进行 HIPAA 标准化、安全性和保健系统集成方面的工作。他在软件开发和设计领域有 18 年以上的经验，持有计算机科学学士学位和 MBA 学位。

技术编辑简介

Patricia Moulder（CISSP、CISM、NSA-IAM）是一位高级安全主题专家和顾问。她持有东卡罗莱纳大学科学硕士学位。她在网络安全评估、Web 应用审计、商用及美国政府客户无线网络技术方面有超过 19 年的经验。她在辛克莱尔社区学院担任网络安全助理教授 5 年之久，她在 SDLC 应用安全审计和数据隐私标准化方面也有大量跨平台经验。

Preface 序　言

从我第一次遇到 syslog 算起已经有 25 年了。当时我是由围绕一台 Sun-3 新组建的 Sun-2 小型群集系统的管理员，试图通过电话和一位朋友一起调试 UUCP 连接，这位朋友告诉我“检查日志”。观察 syslog 有时令我昏昏欲睡，但我发现，这是查看计算机实际上“做了什么”的最好办法。分屏系统可以在屏幕的上角打开一个窗口运行“ tail–f/usr/spool/messages”命令，这展示了其价值；我可以观察电子邮件的工作、USENET 新闻流、进程启动和停止，我可以看到自己的计算机实际在做什么！直到几年之后我第一次遇到安全事故，我才发现日志对于观察过去和当前的情况都很有用。之后，系统管理员在屏幕的一角观察其日志的时代已经一去不复返，现在，这一切已经很模糊了。

为什么我们当中的有些人痴迷于日志，而其他人则对其漠不关心？我认为这是因为，系统管理员有可能认为“日志记录了某些事情”因而“计算机做了某些事情”以及“一切正常，所以我很高兴”。我的第一个日志分析算法很简单：

```
If the syslog stops, the Pyramid's I/O processor is wedged again
```

如果日志工具能够一次又一次地帮助你摆脱困境，你就会不断地使用它。25 年来，我从 syslog 开始，用日志工具完成了如下工作：

- 搜寻对某个主要电子商务零售商发动“购物车穷举攻击”的痕迹。
- 找出 1000 多个系统中恶意软件自动投放的位置。
- 将我暑假中的一些时光花费在分析一家超级计算机中心 10 年的日志，偶然地发现只通过日志的体积就能够检测主要的 Linux 发行版本。
- 构建一个网站数据复制系统，使用 syslog 作为原子事务记录。
- 确认谁向 President Clinton @whitehouse.gov 发送了一封危险电子邮件。
- 计算我的软件工程师在 Diablo Ⅱ游戏上花了多少时间。
- 解析和重放 6 个月的日志中的事务，重新构造一个损坏的数据库。

除此之外，我还用 syslog 做许多“日常工作”，例如确保系统一切正常，寻找不寻常的活动，以及找出故障点及其原因。和计算机中的其他系统不同，我们可以用系统日志对过

去的工作进行限制性查看，限制来自我们收集的信息，以及保存的时长，这是一个有用的视图。因此我总是说："如果你看到一个系统的日志是关闭的，就说明系统管理员对其工作毫不在乎。"。我不知道曾经有多少次听到在事故响应中，关键的系统"因为性能的原因"关闭了日志。对于我这样痴迷于日志的人来说，那样的做法完全没有经过深思：性能问题很容易用更快的处理器或者一个闪存盘来解决，而没有了日志，你就是在"盲飞"。

奇怪的是，关于系统日志的好书并不多。你可能以为，有许多的书籍能够告诉你有用的日志是什么样的，但是实际上没有，这可能是因为该主题有些枯燥，而且这个主题太过广泛，需要覆盖从获取数据到如何处理所获取数据的所有细节。而且还有一个大问题：如何处理日志本身没有简单的方法。"在星期天需要对 syslog 做的前 10 种处理"这样的简单规则是没有的，因为每个人的日志都不相同，需求也各不相同。写一本从本质上说明"重点，查看你的日志并且思考你所看到的内容"的书籍是很困难的。安全人员想要看的是入侵的企图；系统管理员看到的是表示系统正常运行的特征；CIO 将会查看使用情况指标和业务证据；审计人员将会查看可以选择的复选框，凡此种种。窍门就是向所有这些人解释：系统日志是一种集地板蜡、甜点配料和足底按摩器为一体的全能产品，哦，顺便提一句：它只适用于自己动手解决问题的人。

日志需要自己动手这种特性可能是日志难以令人喜欢的原因。而这本书正是你们所要寻找的日志"菜谱"，但是没有简单的"通用"技巧。你必须思考每一种思路，然后对其进行改编，应用到你的特殊情况中。当我教授系统日志分析课程（很久以前的事了！）时，我总是发现班上的有些学生很不愉快地离开了：他们期待在离开时带走"真正能揭示一切的饼状图"或者"具备强大查找功能的日志分析 Perl 脚本"。与此相反，我讲的是通过数据思考分析和分类的框架。我曾经在课程的先决条件中要求"必须了解某种编程方法"，还记得有一位学生最喜欢的日志分析编程语言是 MATLAB。不管你得到什么数据，你所要寻找的就是工作的最佳工具。

在序言中加入我自己的建议可能不太合适，但是无论如何我还是要这么做。对日志进行有效操作的最佳途径是召集 3 ~ 4 位能干的人一起开会，买一些披萨和啤酒，花上两个小时审视你的日志。将日志的内容投影到大屏幕上，让每个人都能看见，并且来回滚动，查看你所得到的结果。然后，当你觉得厌倦的时候开始问自己：你从日志中得到了哪些想要知道的信息，你想要生成哪些方面的摘要，哪些内容能够组成系统工作完成的有用指标，哪些内容可能指出关键的错误。本书对于如何做到这些提供了更多有用的细节，但是相信我，买些披萨和啤酒是没错的。

我可能已经说了太多你已经知道的事情，所以该到此为止了。现在，翻过这一页，开始阅读本书吧！

Marcus J.Ranum

Tenable Network Security 公司 CSO

Preface 前　言

欢迎阅读本书。本书的目标是向信息技术（IT）专业人士提供理解和处理日志数据的入门知识。各种形式的日志数据是由许多类型的系统生成的。如何处理和分析日志数据是长期存在的一个问题。本书介绍能够帮助你分析日志数据和寻找恶意活动的技术和工具。

过去，系统管理员审阅日志文件，寻找磁盘错误或者内核问题。现在的系统管理员往往还要兼任安全管理员。更好地理解如何处理安全日志数据的需求从未像今天那么重要。安全性分析人员是IT专家组中负责跟踪日志分析技术的人。许多经验丰富的人曾经在“压力测试”的模式下进行学习。本书的目标是提供帮助你快速理解各种概念的材料，提取许多人曾经花费多年学习的知识。

我们先来谈谈最近的一个重要问题：法规依从性。随着安能公司和其他企业的垮台，法规依从性现在成为了许多企业的核心主题，焦点现在集中在政策和规程。作为IT工程师，你能说明，Bob在被解雇之后无法访问自己的企业电子邮件账户吗？这些事情需要公司提出规程。系统和网络日志的框架正在以这样那样的方式变化。

目标读者

本书面向任何对学习日志记录和日志管理感兴趣的人。下面介绍一些应该阅读本书的人士。

系统管理员：承担企业日志数据监控任务。

初级安全工程师：网络安全的新手，希望学习日志分析技术。

应用程序开发人员：对从头构建一个日志分析系统感兴趣的人，本书提供了这方面的示例代码。但是，全书为日志分析的重要性提供了很好的背景资料，这些领域也不应该忽略。

管理人员：想深入理解日志数据收集、存储、分析和法规依从性等主题的人。如前所述，这些问题在企业中越来越多地出现，并将持续地成为IT专业人士需要更多关注的领域。

所需基础知识

我们假定读者对网络、操作系统和网络安全等概念有基本的了解。但是，理解本书中的素材并不要求读者是计算机科学家或者网络高手。需要背景信息的主题提供了必要的细节。大部分代码示例采用 Perl 和 Java 编程语言。理解或者使用这些代码示例并不要求读者是 Java 高手，所以我鼓励每个人都仔细地阅读这些例子。

本书的组织

本书的每一章内容都建立在前一章的基础上。尽管如此，许多章节可以独立阅读。本书共有 22 章。

第 1 章：木材、树木、森林

该章提供了日志系统的背景信息。如果你熟悉了 syslog、SNMP、安全日志、日志数据收集、存储等概念，就可以跳过本章。

第 2 章：日志是什么

该章描述日志消息，包括日志重要性的讨论。

第 3 章：日志数据来源

该章描述 syslog 协议、SNMP 和 Windows 事件日志。此外，介绍了日志数据源的分类。

第 4 章：日志存储技术

如果你想要学习有关日志保留、存储格式和在关系数据库管理系统（RDBM）中存储日志的知识，那么请详读该章。我们甚至提供了使用 Hadoop 进行这方面工作的例子。

第 5 章：syslog-ng 案例研究

该章深入介绍了如何在实际环境中部署 syslog-ng 进行日志收集，还讨论了 syslog-ng 的一些较为高级的功能。

第 6 章：隐蔽日志

如果你需要以隐蔽的方式使用日志，该章提供了完成这一任务的许多细节。

第 7 章：分析日志的目标、规划和准备

在你开始分析日志数据之前，首先需要设定目标、规划并为任务做准备。该章介绍的主题包括寻找过去的问题、未来的问题和从未见过的情况。

第 8 章：简单分析技术

在讨论高级分析技术之前，需要介绍基础知识，这包括人工日志分析及其工具。除此之外，我们讨论一个可以使 Windows 事件日志的阅读更为简单的高级工具。当然，我们还讨论了响应日志分析结果的过程。

第 9 章：过滤、规范化和关联

该章是激动人心的一章，介绍了相关的技术和工具，能够帮助你执行关联分析，找出简单的人工日志分析可能忽视的问题。该章介绍的主题包括过滤、规范化、分类学、关联和一些常见的搜索模式。该章中有两节是为对构建自己的关联分析引擎感兴趣的开发人员而写的。该章还介绍了 Jess 和 Esper，说明如何构建基于规则和基于流的引擎。

第 10 章：统计分析

该章讨论如何使用统计执行分析，介绍了频率统计、基线、阈值和异常检测。我们甚至提供了使用机器学习进行分析的方法。

第 11 章：日志数据挖掘

该章专门介绍日志挖掘或者日志知识发现——不同类型的日志分析，它不依赖于对所要寻找的内容的了解。这打破了日志搜索中对字符串列表或者模式的依赖，将日志分析的“高级艺术”提高到一个新的水平。

第 12 章：报告和总结

该章关注作为日志分析手段的报告，我们特别重视对日志数据最佳报告的定义。

第 13 章：日志数据可视化

日志数据的可视化往往很有用。我们的意思不是查看警报、电子邮件或者特定日志分析系统的输出。我们更感兴趣的是讨论在有向图和其他可视化工具中查看日志数据。

第 14 章：日志法则和日志错误

该章介绍各类组织在日志上常见的错误（以及正在犯的错误），还介绍一些组织处理日志中常见的规则和相关性，“法则”的叫法可能过于严格了。

第 15 章：日志分析和收集工具

该章提供了用于日志数据分析和收集的开源和商用工具的评论，这些评论可为读者在选择日常数据管理工具时提供一些参考信息，包括使用这些工具进行日志分析的例子，并且包含了使用这些工具审核常见日志任务和场景的实例。该章将帮助读者评估可用的工具集，以便确定自己组织中分析日志的合适工具。

第 16 章：日志管理规程

该章对日志管理中的日志审核、响应和升级做了介绍。使用支付卡行业（Payment Card Industry，PCI）数据安全性标准（Data Security Standard，DSS）的例子是该章的主题。思路是阐述如何在现实中应用概念。这意味着，示例针对 PCI 标准，但是它们很容易经过改编，扩展到任何环境中。本质上，该章开发了你可以使用的一组步骤和规程。额外的一个好处是对日志管理中解读和应用标准的深入了解。

第 17 章：对日志系统的攻击

该章介绍了对日志记录、日志分析系统的攻击，甚至日志分析人员对日志安全、操作和依从性的破坏。

第 18 章：供程序员使用的日志

该章对于各种程序员都是实用的。这包括系统管理员、Perl 程序员、C/C++ 程序员、Java 程序员及其他编程人员。本质上，任何编写脚本、程序或者软件系统的人都将从该章的内容中获益。人们常说，糟糕的日志消息是糟糕的程序员造成的。虽然这不完全对，但该章的目的就是通过向程序员和其他人提供如何生成更好的日志消息的指导和概念，进而改变这一看法。最终，这些方法将对调试、信息收集和解析能力带来帮助，并且增进软件生成的日志消息的整体实用性。

第 19 章：日志和依从性

该章介绍日志记录和法规及政策的依从性，该章对于任何苦于应付法规依从性的人来说都是有价值的。

第 20 章：规划自己的日志分析系统

该章为日志分析系统的部署规划提供实用的指南。该章的目的不是提供安装特定日志分析系统的详细蓝图，而是提供素材，使你能够将概念应用到任何自己所处的日志分析部署中。该章将告诉你在这一过程中所要询问的问题和考虑的项目。

第 21 章：云日志

云计算是现在的热门话题，并将越来越热门。我们目睹了传统的紧缩型套装产品从公司所属的数据中心移植到云中（这一切正在发生），IT 管理员在硬件、交换机、机架、软件上花费较少的资本性支出（CAPEX），控制日志数据收集、集中和存储甚至安全信息及事件管理（SIEM）的机会将会大大减少。该章介绍云计算和云日志，还介绍与云环境相关的法规和安全问题、云中的大数据、云中的 SIEM、优缺点以及几个关键云日志提供商的详细清单。

第 22 章：日志标准和未来的趋势

该章提供了日志标准的未来和日志记录及日志分析未来发展的专业意见。

致谢

Dr.Anton A.Chuvakin

首先，也是最重要的：感谢我的妻子 Olga，她是我所有作品永恒的灵感源泉，感谢她提供的宝贵的项目管理建议，感谢她容忍（是的，几乎总是在容忍……）我将本可以共度的夜晚用在本书上。

接下来，我特别感谢 Marcus Ranum 为本书所作的序。

最后，感谢 Syngress/Elsevier 的全体人员，感谢他们对我们未能按照承诺的时间交稿的宽容。

Kevin J.Schmidt

首先，我要感谢美丽的妻子 Michelle。她给我的鼓励和支持帮助我完成了本书。当然，感谢我的雇主 Dell，为我完成这一项目提供了支持。接下来必须感谢提供宝贵意见的同事们：Rob Scudiere、Wayne Haber、Raj Bandyopadhyay、Emily Friese、Rafael Guerrero-Platero 和 Maro Arguedas。来自 BalaBit IT Security 的 Robert Fekete 在 syslog-ng 的有关章节上提供了杰出的意见。Ernest Friedman-Hill 为第 9 章中关于 Jess 的小节提供了宝贵的建议。我过去的同事 Jimmy Alderson 慷慨地为第 13 章提供了代码示例。最后，我要感谢合著者 Anton 和 Chris，感谢他们为这本出色的书籍提供了杰出的内容。

Christopher Phillips

我要感谢美丽的妻子 Inna 和可爱的孩子们，Jacqueline 和 Josephine。她们的亲切、幽默和爱，在我编写本书和经历生活中所有努力和冒险的时候提供了灵感和支持。我还要感谢我的父亲一如既往地支持和鼓励我从事工程和科技。感谢 Rob Scudiere、Wayne Haber 和我的雇主 Dell 提供的宝贵意见。我还要特别感谢合著者 Kevin 提供机会，让我成为这本出色书籍的作者之一。Kevin 在 Dell SecureWorks 和我共事的许多年中提供了很好的指导和鼓励，帮助我在职业生涯中变得更为专业。他的领导能力和安全知识是我、我们的客户和每天在一起的许多人灵感的源泉。

目　录 *Contents*