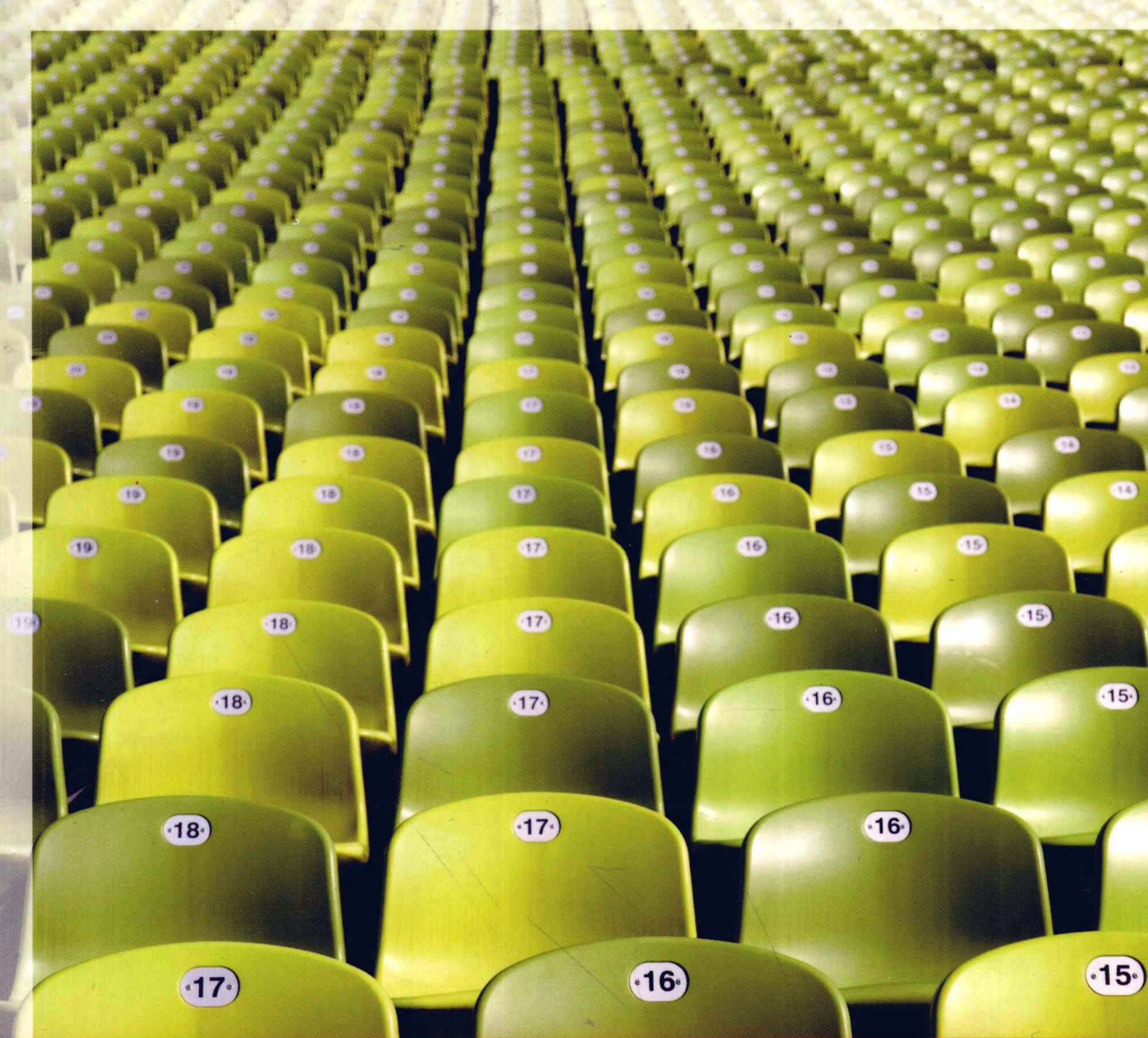




/THEORY/IN/PRACTICE

Web容量规划的艺术

The Art of Capacity Planning

93.092
6

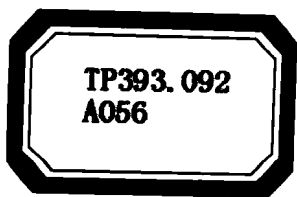
EILLY®



机械工业出版社
China Machine Press



John Allspaw 著
叶飞 罗江华 译



-6

Web容量规划的艺术



John Allspaw 著

叶飞 罗江华 译

TP393.092

A056

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Taipei • Tokyo

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

Web容量规划的艺术/ (美) 阿尔斯帕瓦 (Allspaw, J.) 著; 叶飞、罗江华译. —北京: 机械工业出版社, 2009.8

书名原文: The Art of Capacity Planning

ISBN 978-7-111-27401-8

I. W… II. ①阿… ②叶… III. 主页制作—程序设计 IV. TP393.092

中国版本图书馆CIP数据核字 (2009) 第092367号

北京市版权局著作权合同登记

图字: 01-2009-1578号

©2008 by O'Reilly Media, Inc.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and China Machine Press, 2009.
Authorized translation of the English edition, 2008 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2008。

简体中文版由机械工业出版社出版 2009。英文原版的翻译得到O'Reilly Media, Inc.的授权。此简体中文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。

本书法律顾问

北京市展达律师事务所

书 名/ Web容量规划的艺术

书 号/ ISBN 978-7-111-27401-8

责任编辑/ 陈佳媛

封面设计/ Mark Paglietti, 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/ 北京京北印刷有限公司

开 本/ 178毫米×233毫米 16开本 8.5印张

版 次/ 2010年1月第1版 2010年1月第1次印刷

定 价/ 29.00元 (册)

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

购书热线: (010) 68326294; 88379649; 68995259

投稿热线: (010) 88379604

读者信箱: hzjsj@hzbook.com



Web容量规划的艺术

O'Reilly Media, Inc.介绍

为了满足读者对网络 and 软件技术知识的迫切需求，世界著名计算机图书出版机构 O'Reilly Media, Inc. 授权机械工业出版社，翻译出版一批该公司久负盛名的英文经典技术专著。

O'Reilly Media, Inc. 是世界上在 Unix、X、Internet 和其他开放系统图书领域具有领导地位的出版公司，同时也是联机出版的先锋。

从最畅销的 *The Whole Internet User's Guide & Catalog*（被纽约公共图书馆评为20世纪最重要的50本书之一）到 GNN（最早的Internet门户和商业网站），再到 WebSite（第一个桌面PC的Web服务器软件），O'Reilly Media, Inc. 一直处于Internet发展的最前沿。

许多书店的反馈表明，O'Reilly Media, Inc. 是最稳定的计算机图书出版商——每一本书都一版再版。与大多数计算机图书出版商相比，O'Reilly Media, Inc. 具有深厚的计算机专业背景，这使得 O'Reilly Media, Inc. 形成了一个非常不同于其他出版商的出版方针。O'Reilly Media, Inc. 所有的编辑人员以前都是程序员，或者是顶尖级的技术专家。O'Reilly Media, Inc. 还有许多固定的作者群体——他们本身是相关领域的技术专家、咨询专家，而现在编写著作，O'Reilly Media, Inc. 依靠他们及时地推出图书。因为 O'Reilly Media, Inc. 紧密地与计算机业界联系着，所以 O'Reilly Media, Inc. 知道市场上真正需要什么图书。

作者简介

John Allspaw目前就职于Flickr.com并担任工程经理一职，该网站以共享用户上传的照片闻名。自该网站1999年成立以来，使他聚集了丰富的经验。这些经验包括在线新闻杂志（比如：Salon.com、InfoWorld.com、Macworld.com）以及一些当前急速增长的社会站点（比如：Friendster 和Flickr）。

John在Friendster公司时，该网站曾呈五倍增长。他负责将Friend-ster站点从只有几十台服务器的数据中心过度到多于400台服务器的两个数据中心，以支持重新设计的后端基础设施。当他加盟Flickr公司时，在温哥华只有一个10多台服务器的微小数据中心，现在在美国已经设立了多个数据中心。

在他从事网站职业之前，John曾作为机械工程师在建模和仿真领域为国家公路交通安全局的汽车碰撞模拟实验做出贡献。

封面介绍

我们的封面是根据读者的反馈、自己的试验以及从各分销渠道反馈回来的信息而形成的，颇具特色的封面风格与我们与众不同的技术主题互为补充，使得原本枯燥无味的主題显得鲜活生动。

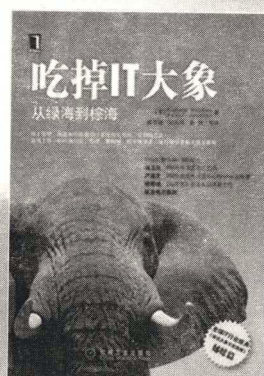
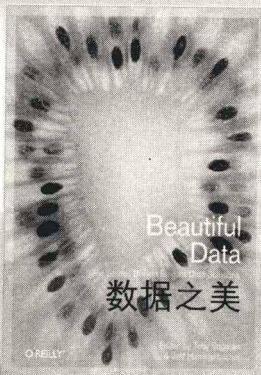
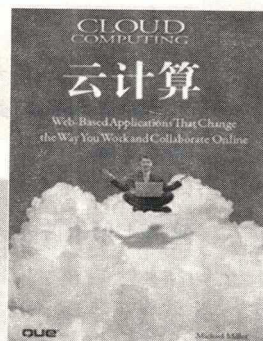
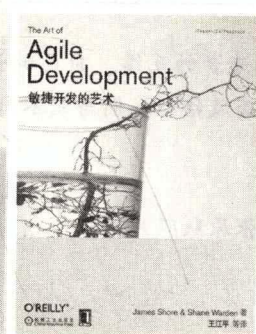
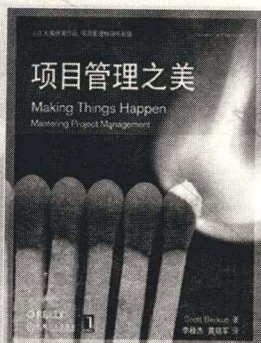
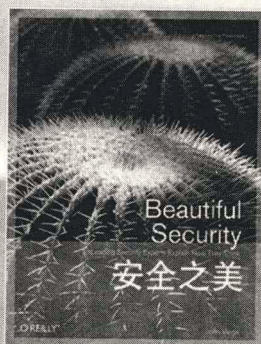
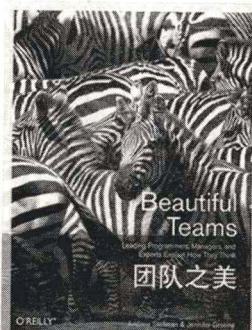
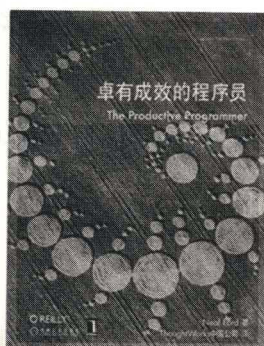
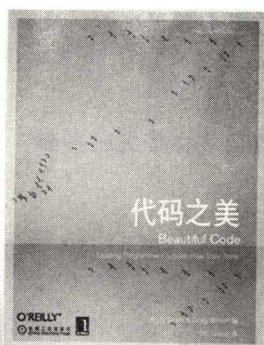
封面上的图片是一张原版照片。



机械工业出版社
CHINA MACHINE PRESS

为每一个团队提供最有价值的阅读服务

每一本书都值得您和您的团队一起阅读
分享阅读 分享成功





专业成就人生
立体服务大众

www.hzbook.com

填写读者调查表 加入华章书友会 获赠精彩技术书 参与活动和抽奖

尊敬的读者：

感谢您选择华章图书。为了聆听您的意见，以便我们能够为您提供更优秀的图书产品，敬请您抽出宝贵的时间填写本表，并按底部的地址邮寄给我们（您也可通过www.hzbook.com填写本表）。您将加入我们的“华章书友会”，及时获得新书资讯，免费参加书友会活动。我们将定期选出若干名热心读者，免费赠送我们出版的图书。请一定填写书名书号并留全您的联系信息，以便我们联络您，谢谢！

书名：

书号：7-111-()

姓名：	性别： ☐ 男 ☐ 女	年龄：	职业：
通信地址：		E-mail：	
电话：	手机：	邮编：	

1. 您是如何获知本书的：

☐ 朋友推荐 ☐ 书店 ☐ 图书目录 ☐ 杂志、报纸、网络等 ☐ 其他

2. 您从哪里购买本书：

☐ 新华书店 ☐ 计算机专业书店 ☐ 网上书店 ☐ 其他

3. 您对本书的评价是：

技术内容	☐ 很好	☐ 一般	☐ 较差	☐ 理由_____
文字质量	☐ 很好	☐ 一般	☐ 较差	☐ 理由_____
版式封面	☐ 很好	☐ 一般	☐ 较差	☐ 理由_____
印装质量	☐ 很好	☐ 一般	☐ 较差	☐ 理由_____
图书定价	☐ 太高	☐ 合适	☐ 较低	☐ 理由_____

4. 您希望我们的图书在哪些方面进行改进？

5. 您最希望我们出版哪方面的图书？如果有英文版请写出书名。

6. 您有没有写作或翻译技术图书的想法？

☐ 是，我的计划是_____ ☐ 否

7. 您希望获取图书信息的形式：

☐ 邮件 ☐ 信函 ☐ 短信 ☐ 其他_____

请寄：北京市西城区百万庄南街1号 机械工业出版社 华章公司 计算机图书策划部收
邮编：100037 电话：(010) 88379512 传真：(010) 68311602 E-mail: hzjsj@hzbook.com

目录

前言	1
第1章 容量规划的目标、问题和过程	7
快捷但不好的数学	8
预测你的系统何时会失败	9
用系统统计表呈现问题	10
买东西：采购是一个过程	11
性能与容量：两种不同的概念	12
社交网站和开放式API的影响	13
第2章 设定容量目标	15
不同种类的需求和测量方法	15
架构决策	19
第3章 测量：容量的单位	27
容量跟踪工具的方方面面	28
应用程序监测	35
API的使用率及其对容量的影响	62
示例和现实	64
小结	65

第4章 趋势预测	66
曲线拟合	67
采购	82
增加容量后的影响.....	85
长期趋势	85
迭代和校准.....	89
小结	92
第5章 部署	93
自动化部署基本原理.....	93
自动化安装工具	96
自动化部署	99
小结	103
附录A 虚拟化和云计算.....	105
附录B 对瞬时增长的处理	119
附录C 容量工具.....	123

前言

2005年7月7日凌晨3点左右，我和我的同事卡尔·亨德森正在处理将网站 Flickr.com的所有访问流量迁移到它的新家（位于德克萨斯州的一个Yahoo! 数据中心）的收尾工作。原来在温哥华的那些基础设施的超负荷现象越来越严重，而且受到电力和空间的严重限制。因为雅虎刚刚收购了Flickr，所以是时候提高它的在线容量了。当我们将DNS记录指向崭新的服务器后，大约过了一个小时，卡尔不经意间看到了一则新闻：伦敦地铁刚刚遭遇了炸弹袭击。

伦敦市民用具有拍照功能的手机和其他设备记录下了发生的一切。在接下来的24小时里，Flickr的访问流量比以往任何时候都大，因为来自灾难现场的照片被不断地上传到网站上面。新闻也开始链接到这些照片，新服务器的访问流量因此到达了峰值。

这不只是全民从事新闻工作的一个极佳范例，也是网站容量规划的一堂实物教学课。不幸的是，它来自于一场灾难。网络访问流量其有偶然性和不可预测性。如果我们没有将Flickr及时迁移到新的数据中心，它那天也许会宕掉。

容量规划古已有之，从经济学到工程学等领域都有所应用。通俗地讲，容量规划就是资源管理。当资源有限且具有一定成本时，你就需要进行容量规划。

当一家土木工程公司设计一个新的高速公路系统时，它需要对车辆承载容量进行规划，正如一个为大城市提供电力的能源公司需要进行容量规划一样。在某些方面，他们关注的重点和网站运营有很多共同点，许多基本的概念和重点都可以应用于这三门学科。

虽然系统管理在20世纪60年代就已经存在，但是专注于为Web站点提供服务的分支还是新出现的。网站运营的很大一部分工作就是站点容量的规划和管理。这些只是过程而不是目标，并且它们由不同的部分组成。尽管每个组织采取的方式各不相同，但是基本原理还是一样的。

- 确保有合适的资源（服务器、储存设备、网络等）能用于处理预期的或突发的负载。
- 有明确的采购和审批机制。
- 准备好调整资金支出，以支持业务需求。
- 有适当的部署和管理系统用于管理部署后的资源。

我为什么写这本书

我担任运营工程经理时一直面临着一个难题，那就是无法从任何途径获取到关于在同一时刻该让多少设备保持运行状态的帮助信息。现有书籍中关于计算机容量规划的内容更偏重于数学理论上的资源规划，而不是整个过程的实际执行。

许多著作只讲解了Web站点用例的基本模型，缺乏详细而明确的信息或建议。相反，它们只提供了被设计用来阐述排队理论原理的数学模型，这些数学模型是传统容量规划的基础。这种方法也许在数学上是很有趣和一流的，但是当运营工程师被告知只有一周的时间用来为不可预知的额外流量（可能是由于某一项超棒的新功能发布带来的）做准备时，或者是当他看到自己的网站因Yahoo!，Digg或CNN等的首页链接带来的负荷而即将崩溃时，这种方法对他们并没有帮助。

我发现大部分有关网站容量规划的书都暗含着这样的假设：在非网络环境（比如制造业或工业工程）中建立的过程和概念在网络环境中也同样适用。尽管与规化等相关的理论实际上是相似的，但是这些概念的实际应用并不能与网站的短期发展很好的对应。

据我观察，大多数网络开发环境都变化得太快太频繁，以致于不能像其他领域一样，可以进行详细且严格的容量调研工作。当运营工程师在系统中提出排队模型时，新的代码已经部署完毕，并且使用的特点有可能已经发生了巨大变化。又或者是，一些在科技、社会或现实世界发生的事情使得所有的建模与仿真变得过时了。

与工业界同事的探讨让我受益良多，因为他们也同样遇到了很多关于可扩展性和容量方面的问题。随着时间的推移，我接触过许多不同的公司，每家公司采用不同的架构，每一种架构都遇到不同的问题。但是，他们采用的解决方案非常相似，我希望我能在本书中阐明这些解决方案。

本书的重点和主题

本书不是关于如何构建复杂的模型和仿真的，也不是关于耗费大量时间来反复运行基准测试的。同时，本书也不是关于利托氏定理、马尔可夫链（译注1）或泊松到达率这样的数学原理的。

本书将介绍如何在现实世界中切实有效地执行容量规划和管理。本书还会介绍实际工具的使用，以及如何改变这些工具的用途以使它适应网站的发展。当你在高速公路上遇到爆胎时，你可能会花大量时间来努力找出爆胎的原因，也可能会直接安装好备用胎，然后重新上路。

这就是我正在采用的容量规划方法：自适应法，这并不是理论上的。

请记住，本书中有大量的信息看起来像是常识——这是一件好事。解决问题的最简单的方法往往也是最好的，容量规划也不例外。

本书将讲解对那些不断发展的网站进行容量规划的整个过程，包括测量、采购和部署。我将介绍一些久经考验且受欢迎的测量工具和技术。这些工具大多数都可以同时在LAMP（译注2）环境和基于Windows的环境下运行。因此，我在介绍时会尽可能采用与平台无关的方式。

译注1： 利托氏定理（Little's Law）：

在一个稳定的系统中，平均的客户数（访问该系统的客户）等于客户到达该系统的平均速率乘以客户在该系统中花费的平均时间，这是一个粗略的估算公式。但是，它对估算开发的系统的性能非常有用。特别是当我们开发服务器端的应用时，用这个公式来估算系统所承受的压力是非常有用的。所谓稳定的系统，指的是“客户到达该系统的平均速率恒等于客户离开该系统的平均速率”。

马尔可夫链（Markov chains）：

因安德烈·马尔可夫（A.A.Markov, 1856-1922）得名，是数学中具有马尔可夫性质的离散时间随机过程。该过程中，在给定当前知识或信息的情况下，过去（即当期以前的历史状态）对于预测将来（即当期以后的未来状态）是无关的。

译注2： LAMP是指Linux+Apache+MySQL+Perl/PHP/Python，一组常用来搭建动态网站或者服务器的开源软件，本身都是各自独立的程序，但是因为常被放在一起使用，拥有了越来越高的兼容度，共同组成了一个强大的Web应用程序平台。

当然，数据库、Web服务、缓存服务和存储解决方案等方面的细节内容已经超出了本书的范围。我将利用例子来阐述每一个相关的过程和概念，但本书并不是一本实施指南。这样做是为了能够用尽可能通用的方式来阐述资源管理——资源管理过程本身也是我们希望重点强调的。

例如，数据库用于存储数据和对查询进行响应。市面上大多数流行的数据库都允许将数据复制到其他服务器中，从而增强它们在冗余性、性能和架构方面的竞争优势。这样还有助于在Postgres、Oracle或MySQL之间实现数据复制这样的技术（在别的书里有相关的主题）。本书讲解的是那些与容量规划和部署有关的数据复制技术。

实际上，这本书是关于如何测量、规划和管理正在发展的网站应用的，与你选择的基础技术无关。

本书为谁而写

本书适合于系统、存储、数据库和网络等领域的管理员、工程经理和容量规划师。

本书适合于那些希望（也许害怕）自己的网站将来成长为像Facebook、Flickr、Myspace、Twitter和其他一些经历过因访问量激增而进行大规模扩容的公司。本书中讲述的方法都来自于真实网站的经验，这些网站都经历了从流量大增到迅速壮大这一过程。如果你也期望你的网站流量大增，那么请阅读本书。

本书是如何组织的

第1章：容量规划的目标、问题和过程。本章介绍了一些访问量非常大的一些站点会经常遇到的问题。

第2章：设定容量目标。本章阐述了对一个快速增长的Web应用进行容量规划时会涉及的各种问题，以及规划的容量如何适合于整个站点的可用性和性能。

第3章：测量——容量单位。本章讨论了容量的测量和监控。

第4章：趋势预测。本章解释了如何将测量数据转换成预测，以及趋势分析是如何适合整个规划过程的。

第5章：部署。本章讨论了与部署相关的概念：自动化安装、配置和管理。

附录A：虚拟化和云计算。本章讨论了怎样在容量规划中运用虚拟化和云服务。

附录B：对瞬时增长的处理。本章为陷入容量危机的站点提供了建议，同时也为处理站点故障提供了最佳实践。

附录C：容量工具。这是一份关于测量、安装、配置和管理工具的说明清单，这些工具贯穿全书并高亮显示。

本书采用的约定

斜体 (*Italic*)

对新术语的声明、URL、文件名、UNIX工具、操作语言、命令行选项。

等宽字体 (`Constant width`)

表示文件内容、命令输出和常规代码。

等宽粗体 (**`Constant width bold`**)

表示命令或其他一些必须用户手动输入的文档、部分代码或用于表示讨论的文件标注。

等宽斜体 (*`Constant width italic`*)

表示应该由用户自定义的值来代替的文字。

使用示例代码

这是一本能帮你顺利完成工作的图书。实际上，你可以在你的程序或文档里使用本书附带的代码。你不必联系我们获得许可，除非你是要复制一些重要的代码。比如说，写的某个程序用到了大量本书的代码、出售或者分发O'Reilly的书赠CD光盘而没有寻求许可。引用本书的示例代码回答提问也不需要许可。对本书中的部分示例代码进行重新整合进你的产品文档不需要许可。

我们真心希望，但是不强迫加上引用标题。一个引用标题通常包括书名、作者、出版商和ISBN。比如：“The Art of Capacity Planning by John Allspaw Copyright 2008 Yahoo! Inc., 978-0-596-51857-8.”

如果你发现在使用代码示例过程中有问题或者需要获得其他的使用许可，请联系我们 permissions@oreilly.com。

联系方法

如果您对本书的意见和问题，请联系出版社：

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

我们对本书有专门的网页来提供额外的代码和更多的附加信息，你可以访问这个页面地址：

<http://www.oreilly.com/catalog/9780596518578>。

提错或者有问题请发邮件到：

bookquestions@oreilly.com。

关于我们的书籍、会议、资源中心、O'Reilly网络的更多信息，请访问我们的站点：

<http://www.oreilly.com>

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

致谢

在这样简短的篇幅里，寥寥数语不足以感谢每个人，不过我会尽量提到他们的名字。本书中的大多数素材源自他们的实践经验，很多人陪我坚持到了最后。特别感谢Peter Norby、Gil Raphaelli、Kevin Collins、Dathan Pattishall、Cal Henderson、Aaron Cope、Paul Hammond、Paul Lloyd、Serguei Mourachov 和Chad Dickerson。同样感谢Heather Champ 和整个Flickr 客服团队。感谢Flickr的开发工程师，你们都像运营工程师一样深思熟虑，这让我很感激。多亏Stewart Butterfield和Caterina Fake在初期说服我加入Flickr团队，也多亏David Filo和Hugo Gunnarsen促使我用真实数据备份硬件需求。感谢Kevin Murphy，在自动化部署章节提供了很多的素材。感谢Andy Oran和Isabel Kunkle的编校， 特别感谢我的好朋友Chris Colin的优秀预编建议。

感谢Adam Jacob、 Matt St. Onge、Jeremy Zawodny和Theo Schlossnagle的技术评审。

非常感谢Matt Mullenweg 和Don MacAskill共享他们关于云基础设施的使用案例。

最重要的是要感谢我的妻子Elizabeth Kairys，她给了我最大的鼓励和支持。没有她，这本书根本无法完成。

容量规划的目标、 问题和过程

设计本章的目的在于帮助你装配并运用后面章节中将介绍的大量工具和技术。如果你没有掌握本章中所介绍的概念，那么当你阅读本书的其余部分时，你就像在不知道如何使用指南针、六分仪或GPS设备的情况下置身于海洋上，永远在原地兜圈子。

如果将容量规划和管理（组织网站正常运转所需的资源应采取的步骤）分解开来，其实过程十分简单。开始之前，你应该问自己一个问题：你的网站需要什么样的性能？

首先，需要使用诸如响应时间、可供消耗容量以及峰值驱动处理等明确的指标来定义应用程序的总体负载和容量需求。峰值驱动处理是指站点访问量处于最高峰时应用程序的资源（服务器和数据库等）所承受的工作负荷。图1-1演示了这一过程并回答了下面这些问题。

1. 当前的基础设施工作状况如何？

测量构成应用程序的每一部分的架构（Web服务器、数据库服务器、网络等）的工作负荷特征，并将它们与你上面提到的性能需求相比较。

2. 你未来需要什么来维持这种可接受的性能？

在你过去对系统的性能的认知的基础上对未来进行预测，然后结合一项你能够承担得起的预测，以及一份切合实际的时间表。确定你需要什么以及何时需要它。

3. 在你收集到你所需要的东西以后，如何安装和管理这些资源？

利用经过工业认证的工具和技术来部署新的容量。

4. 不断地整理、重复、迭代和校准你的容量规划。