

Python Linux

系统管理与自动化运维

赖明星 著

Python for Linux System Administration
and Automated Operation

- 作者就职于腾讯，曾就职于网易，平台开发工程师、数据库内核工程师和高级运维工程师，善于用Python解决系统管理问题
- 从工具、方法、实战三个维度讲解了利用Python进行Linux系统管理和自动化运维的最佳实践



Python Linux

系统管理与自动化运维

Python for Linux System Administration
and Automated Operation

赖明星 著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

Python Linux 系统管理与自动化运维 / 赖明星著. —北京: 机械工业出版社, 2017.9
(Linux/Unix 技术丛书)

ISBN 978-7-111-57865-9

I.P… II. 赖… III. Linux 操作系统 IV. TP316.85

中国版本图书馆 CIP 数据核字 (2017) 第 213433 号

Python Linux 系统管理与自动化运维

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: 何欣阳

责任校对: 李秋荣

印 刷: 北京诚信伟业印刷有限公司

版 次: 2017 年 9 月第 1 版第 1 次印刷

开 本: 186mm × 240mm 1/16

印 张: 27

书 号: ISBN 978-7-111-57865-9

定 价: 89.00 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) 88379426 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzit@hzbook.com

版权所有 · 侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

内容简介

Python是系统管理和运维领域的一把利器，本书是作者多年来在网易的云平台开发、数据库内核开发、系统管理与运维领域使用Python的经验总结。

本书以Linux系统管理为线索，以Python语言为载体，从工具、方法、实战等多个方面讲解了如何在Linux系统管理和自动化运维中使用Python解决各种问题，包含大量案例和最佳实践。

全书逻辑上分为5个部分：

介绍篇（第1章）：介绍了Python语言的优缺点和Python在Linux系统管理领域的应用；

工具篇（第2章）：讲解了多个Python生态工具，充分利用这些工具，不但可以有效地提高工作效率，而且还能形成统一的代码风格；

脚本篇（第3~7章）：详细讲解了如何使用Python编写脚本管理Linux，包括如何使用Python构建命令行工具、如何处理文本、如何进行系统管理、如何监控Linux系统，以及如何处理文档与报告，包含大量的Python实战案例。

自动化篇（第8~10章）：主要讲解了多个Python语言开发工具，包括网络嗅探工具Scapy、自动化运维工具Fabric、自动部署工具Ansible等，灵活使用这些工具，可以有效地提高运维工程师的工作效率。

综合案例篇（第11章）：介绍了一个综合案例，即使用Python打造MySQL数据库专家系统。其中详细讲解了Python中的高级语言特性和系统架构，充分理解这一章的内容，相信读者的Python水平能有一个较大的提升。

作者简介

赖明星 资深云平台开发工程师、数据库内核开发工程师和高级运维工程师，现就职于腾讯CDG（企业发展事业群），曾就职于网易，网易云数据库MySQL云平台、MongoDB云平台的核心开发人员，网易云数据库大规模MySQL系统运维的负责人。

国内知名的Python技术专家和Python技术的积极推广者，知乎专栏《Python高手之路》的所有者和维护者，在网络上撰写了大量优质的原创文章，进行了多次高质量的技术分享。

除Python外，作者对Linux和MySQL也有深入研究，多次在 Oracle 技术嘉年华、IMG（Inside MySQL Group）技术沙龙、全球敏捷运维峰会、PHPCon China等知名技术大会分享MySQL最佳实践。

作者还是IMG社区核心成员，IMG社区自成立以来，努力打造纯净的技术氛围，积极推动技术分享，汇集了各个企业的技术大牛，在技术社区中有着很大的影响力。

扫描下方二维码加入IMG社区



HZBOOKS | 华章科技 | Science & Technology



Foreword 推荐序一

Python 是这个时代最流行的编程语言，不管你承认与否，Python 已然被广泛应用到各个领域。从 Web 开发，到运维开发、搜索引擎，再到机器学习，甚至到游戏开发，没有什么 Python 语言所不能的。特别是在当前这个云计算、大数据、机器学习、AI 蓬勃发展的新时代，Python 正扮演着越来越重要的角色。甚至可以说，作为一名 IT 从业人员，不论你是开发人员、DBA，抑或是前端开发人员，不会 Python 就很可能被这个时代所抛弃。

作者赖明星是我在网易时期的同事兼好友，同事们都戏称其为“网易最牛 Python 程序员”。他不仅有着极其专业的开发能力与丰富的运维经验，更令我佩服和尊敬的是其充满传奇色彩的人生经历，以及在这种人生经历中的不断成长与永不放弃的坚毅精神。对此感兴趣的读者可以看其知乎上的文章与回复。

作者是 IMG (Inside MySQL Group) 社区的成员之一，对 MySQL 数据库有着深刻的理解。同时，作者多次在数据库大会上分享其在数据库方面的工作经验。IMG 社区是一个专注于开源数据库技术，连接一切有态度的技术人员的社区。欢迎感兴趣的小伙伴加入这个社区，微信公众账号：InsideMySQL。若要加入微信群，请先加我个人微信号：82946772。

本书结合了作者丰富的实际工作案例，对 Python 语言做了大量的介绍，可以说是成为 Pythonic 最好的书籍。文章最后的 MySQL 专家系统亦是一个集大成的运维开发系统，同学们可以通过这个案例更好地理解 Python 在运维中的实际应用。最后，我将此书推荐给所有技术圈的同仁们，也希望大家早日体会到 Pythonic 的最高奥义。

姜承尧

腾讯技术总监

推荐序二 *Foreword*

2010年前后，人类迎来了以云计算、大数据、物联网的普及为标志的第三次信息化浪潮，各行各业都在发生着深刻的变革。大数据成为企业提升核心竞争力的关键支撑，越来越多的企业将形成以数据为中心的运营模式，企业对大数据人才的需求正在迅速增长。

数据的采集预处理、存储管理、分析挖掘，是企业中的数据科学家和数据工程师们的核心工作，为了高效地完成上述工作，就需要IT技术人员掌握一门得心应手的编程语言，而Python就是能够满足这种需求的编程语言。Python语言自20世纪90年代初诞生至今，已被逐渐广泛应用于系统管理任务的处理、Web编程和数据挖掘等领域，是目前最受欢迎的编程语言之一。目前，国内的很多高校都在本科阶段开设了Python课程，高校的很多科研群体也都采用Python完成科学数据的分析工作，于是，掌握基本Python编程能力的高校毕业生将不断迈入工作岗位，成为企业中数据分析人才的中坚力量，Python在企业中的使用群体规模会不断扩大。此外，目前市场上流行的大数据处理框架（比如Spark），都支持采用Python进行程序开发，这使得学习者在掌握Python语言以后，能够较好地承担起企业大数据产品开发和系统运维等方面的工作。因此，无论在高校还是在企业，Python语言都已经形成了持续稳定的影响力，Python是学习者在编程语言方面的理想选择。

作者于2011年到2014年在厦门大学数据库实验室攻读硕士学位。在这三年时间里，实验室师生在大数据技术方面进行了系统的学习，以此为基础，实验室编著出版了国内高校第一本系统性介绍大数据知识的专业教材，建设了国内高校首个大数据课程公共服务平台，成为了国内高校大数据教学的知名品牌。在这个过程中，作者在大数据技术研究和学习方面，做出了巨大的努力，尤其是在Python语言方面有了初步的积累。作为导师，令我欣慰的是，作者在步入工作岗位以后，继续发扬了刻苦钻研的精神，不断提升Python语言的实践能力和水平，用Python语言解决了工作岗位中的大量业务问题。最终，作者根据自己丰富的行业实践经验，整理编写了这样一本具有较高实用价值的Python教程。

本书非常适合对一线互联网技术感兴趣的工程师、想要学习 Python 在 Linux 系统管理中如何应用的运维工程师，以及正在寻找 Python 进阶教程的 Python 爱好者进行阅读。相信这本书会成为读者们的良师益友。

林子雨

厦门大学计算机科学系教师

推荐序三 *Foreword*

网易数据科学中心长期专注于数据库系统、搜索系统、大数据计算、BI 等方向，致力于为整个网易集团提供领先的数据基础设施和技术解决方案。在部门十多年的发展过程中，我们一般使用两种语言，十年前我们用 C++ 和 Python，如今我们主要使用 Java 和 Python。Java/C++ 是我们核心系统的开发语言，这两种语言都是当今最流行的编程语言，只有使用最流行的语言才能方便人员招聘、团队协作，从而达到长期维护的目的。随着 IT 行业的发展，Java 替代 C++ 成为了最流行语言，也成为了我们核心系统的开发语言，这是因为当前这个时代更加注重开发和维护效率，而非代码执行效率。Python 语言则应用在 Linux 系统管理、MySQL 系统管理等系统管理场景，Python 简单优美、功能强大，在这些场景下可以秒杀 Java/C++，因此始终是我们必学的编程语言。

市面上有不少讲解 Python 编程语言的书籍，也有不少关于 Linux 系统管理的书籍，但是介绍如何利用 Python 解决 Linux 系统管理问题的书籍却寥寥无几。作者赖明星是网易数据库专家，是数据库云服务（RDS）的核心开发人员。RDS 产品使用 Python 语言成功管理了成千上万个 Linux 主机和 MySQL 实例，成功支撑了考拉、网易云音乐、网易新闻等大型互联网产品，在 RDS 研发运营过程中，作者积累了大量的一线工作经验和案例，并通过这本书分享给大家，希望对大家有所帮助。

余利华

· 网易技术总监

为何写作本书

我第一次接触 Python 是在 2008 年的时候，当时还是一名大学生。有一位非常优秀的学长通过《黑客与画家》这本书了解到 Python 语言，并且认为 Python 将在未来几年迅猛发展，在各个应用领域都会大有可为。因此，在国内还没有广泛使用 Python 语言时，这位学长便推荐我学习 Python。现在的 Python 语言，流行程度已不用我多言，这一方面证实了 Paul Graham（《黑客与画家》的作者）的远见卓识，同时，也反映出这位优秀学长的聪明睿智。从这件事中，我深切地感受到要多与人交流，尤其是要与那些比自己优秀的人交流。正是出于和大家交流的想法，我将这些年的 Python 使用心得，以及在网易云开发与运维的经验汇集成一本书，也就是你现在看到的这本书。

如果大家身处互联网，一定能够切身地感受到 Python 语言的流行程度。Python 语言由于其简单易学、语法优美、应用领域广泛等诸多优点，俘获了大批的粉丝。通过 Google 搜索指数可以看到，Python 语言在最近两年出现了爆发式的增长，其在编程语言排行榜上的排名，从第 8 名一跃上升到第 4 名，并且有继续上升的趋势。越来越多的公司高薪招聘 Python 工程师，越来越多的世界名校开始讲授 Python 语言，越来越多的项目使用 Python 语言进行开发。因此，从各个角度来说，Python 都将大有可为，学习 Python 将是一个正确的选择。

伴随着云计算的高速发展，Python 在系统管理领域也表现不俗。著名的云计算平台 OpenStack 就是使用 Python 语言开发的，各大云计算厂商也选择使用 Python 开发自己的内部系统。并且，由于云计算厂商托管的都是大规模服务器，为了提高运维效率、节约人力成本、快速响应需求，各大云计算厂商都不约而同地使用 Python 进行大规模服务器的部署和运维。

Python 在 Linux 系统管理中，已经担任了非常重要的角色。但是，讲解 Python 在 Linux 系统管理中应用的书籍却寥寥无几，更多都是专注于 Python 语言、Python 爬虫和 Python Web

开发。难得的几本也仅仅是介绍 Python 相关工具的使用而已。由于互联网业务的复杂性和多样性，这远不能满足互联网公司的需求，公司更希望工程师能够灵活地使用 Python，根据业务需求开发自己的工具和系统。因此，本书不仅会介绍 Python 语言编写的系统管理工具，还会讲解如何使用 Python 构建自己的系统管理工具。

主要内容特色

本书介绍了 Python 语言在 Linux 系统管理中的应用，包括编写 Python 脚本管理 Linux 系统，使用 Python 编写的自动化工具管理 Linux 系统，以及使用 Python 打造专属的管理工具等。

本书以 Linux 系统管理为线索，以 Python 语言为载体，介绍了大量 Python 语言的应用案例。通过学习本书，不但可以使用 Python 语言管理 Linux 系统，也为 Python 解决其他领域问题打下了坚实的基础。书中每一章都会解决某一类问题，并且提供了问题的答案。如 Python 管理 Linux 文件和目录，使用 Python 监控 Linux 系统，使用 Python 编写自动化工具，使用 Python 进行自动化运维等。

通过学习本书，你可以：

- 掌握 Python 生态工具，提高自己的开发效率；
- 学会如何使用 Python 构建自己的命令行工具；
- 用 Python 编写可维护性更强的文本处理程序；
- 全方位监控 Linux 系统；
- 使用 Python 自动化部署应用；
- 管理操作系统配置；
- 使用 Python 打造 MySQL 专家系统；
- 通过 Python 发送电子邮件。

本书读者对象

本书不是一本讲解 Python 编程语言的书籍，也不是教授如何使用 Python 运维工具的书籍，而是一本讲解 Python 在 Linux 系统管理中应用的实战书籍。如果你还没学过 Python 编程语言，建议先学习 Python 语言以后，再来阅读本书。如果你从来没有使用过 Linux，书中部分章节可能会让你感到困惑，不过依然有很多章节可以参考。因此，想要更好地学习本书内容，读者需要具备以下条件：

- Python 语言基础知识；
- Linux 使用经验；
- 了解 SSH 的使用。

本书虽然主要讲解 Python 在 Linux 系统管理中的应用，但是，书中很多例子都具有更加广泛的应用场景。本书的最大特色是提供了大量的实战案例，无论是刚学完 Python 语言，正在寻找 Python 实战案例的读者，还是想要学习 Python 在 Linux 系统管理中如何应用的读者，都能从中受益。

综上所述，本书适合以下几类读者：

- 使用 Python 语言的运维工程师；
- 想要提高 Python 技能的开发工程师；
- 想要了解 Python 在互联网应用的在校学生；
- 所有对 Python 实战感兴趣的读者。

如何阅读本书

本书共分 11 章，每一章都可以单独成册。你可以从头开始阅读，也可以选择自己感兴趣的章节阅读。

- 第 1 章介绍了 Python 语言的优缺点和 Python 在 Linux 系统管理领域的应用，这一章主要回答“为什么学习 Python”这个问题。
- 第 2 章介绍了多个 Python 生态工具。充分利用这些工具，不但可以有效提高工作效率，还能形成统一的代码风格。此外，本章还会介绍如何解决 Python 中的环境依赖问题，包括如何在一台服务器上使用不同的 Python 版本，如何对不同的项目安装不同的依赖而不相互影响。相信通过学习本章介绍的工具，能够帮助你解决在学习 Python 过程中遇到的环境问题。
- 第 3 ~ 7 章主要讲解使用 Python 编写脚本管理 Linux，包括如何使用 Python 构建命令行工具、如何使用 Python 处理文本问题、如何使用 Python 进行系统管理、如何使用 Python 监控 Linux 系统，以及如何使用 Python 处理文档与报告。这几章包含了大量的 Python 实战案例，如果你学习本书是为了提高自己的 Python 编程技能，可以重点学习这几章。
- 第 8 ~ 10 章主要介绍了多个 Python 语言开发的工具，包括网络嗅探工具 Scapy、自动化部署工具 Fabric 和自动部署工具 Ansible，这几章较偏重运维操作。灵活使用这几章

介绍的工具有效提高运维工程师的工作效率。Python 语言之所以在系统管理和自动化运维领域使用广泛，是因为有很多使用 Python 语言开发的开源工具。这些工具简单易用、功能强大、扩展性强，是 Linux 系统管理员的得力助手。

- 第 11 章介绍了一个综合案例，即使用 Python 打造 MySQL 数据库专家系统。这一章颇有难度。书中使用数据库专家系统为载体，详细介绍了 Python 中的高级语言特性和 Python 中的系统架构。这一章不仅会用到比较深入的 MySQL 知识，还会用到 Python 的高级技巧。为了打造一个可读性好、可扩展性强的程序，我们使用了不少 Python 语言的高级特性。充分理解这一章介绍的 Python 语言高级特性，相信你的 Python 水平会有一个较大的提升。

需要强调的是，本书是一本实战类的书籍，因此，更多是强调实践的重要性。许多工程师学习 Python，学完就忘，忘了再学，循环往复，都是因为实践不充分导致的。本书包含了大量的实战案例，可以作为 Python 的进阶教材，帮助你举一反三，编写自己的管理程序。

勘误和资源

为了便于实践，书中所有的代码都能够在 GitHub 上找到（https://github.com/lalor/python_for_linux_system_administration.git）。你可以在 GitHub 上查看代码，也可以下载到本地运行。

由于水平有限，加上时间仓促，书中难免出现一些不准确甚至错误的地方，恳请读者批评指正。在此，欢迎大家指出书中存在的问题，并提供指导性的意见，不甚感谢。如果你有任何与本书相关的内容需要与我探讨，都可以来本书源代码的 Issue 页面提问，也可以发邮件到 me@mingxinglai.com，或者在知乎 @mingxinglai 与我联系，我都会及时回复。最后，衷心地希望本书能给大家带来帮助，并祝大家阅读愉快。

致谢

首先，我要感谢厦门大学林子雨老师和腾讯技术总监姜承尧老师。是林老师手把手指导我写论文，才让我具备了写作本书的能力；是姜老师的鼓励和支持，才让我有了写这本书的想法。没有你们的帮助和鼓励，是不可能这本书的。此外，还要感谢网易技术总监余利华老师，是你构建了这个开放的工作环境，才让我有时间和精力来完成这本书。当然，还要感谢三位老师为本书作序。你们是我的人生榜样和奋斗目标，激励着我不断学习和进步，提醒自己时刻不能松懈。

其次，要感谢网易公司的各位同事们，尤其是 RDS 项目组和 TNT 项目组的同事们，能够与你们这样一群才华横溢、激情澎湃的同事一起工作，我感到非常的荣幸和兴奋。也因为这个开放的工作环境中，我才可以不断进行研究和创新。感谢已经离职了的争神，你的一小部分优秀代码出现在了本书之中。还要感谢曾经一起讨论 Python 的小伙伴们，包括章洋、王中、李浩成、王新然、毛茂林、翟亮、陈王仡、刘刚、龚俊、戴晨、黄韬、孙旭东、刘云、钱毅、杨浩、李成、宗雅洁、舒雄、彭应龙等。

我还要感谢我的父亲和母亲，虽然他们在本书的写作过程中没有进行直接的帮助，但是，无论我做什么事情他们都无条件支持我，给予了我无条件的爱。我爱你们。

最后，一份特别的感谢要送给本书的策划编辑杨福川老师，你的专业水平和工作态度，给我留下了深刻印象。还要感谢编辑李艺老师，希望我的写作水平没有让你吃太多的苦。此外，还要感谢出版社其他默默工作的同事们。

目 录 Contents

推荐序一	2.3.1 编写 Python 的 vim 插件····· 15
推荐序二	2.3.2 Windows 下 Python 编辑器 PyCharm 介绍····· 17
推荐序三	2.4 Python 编程辅助工具····· 18
前 言	2.4.1 Python 交互式编程····· 18
第1章 Python语言与Linux系统管理 ··· 1	2.4.2 使用 IPython 交互式编程····· 20
1.1 Python 语言有多流行····· 1	2.4.3 jupyter 的使用····· 29
1.2 Python 语言为什么流行····· 3	2.5 Python 调试器····· 31
1.3 Python 语言有什么缺点····· 4	2.5.1 标准库的 pdb····· 32
1.4 Python 语言的应用场景····· 4	2.5.2 开源的 ipdb····· 34
1.5 为什么 Python 适合 Linux 系统 管理····· 5	2.6 Python 代码规范检查····· 34
1.6 使用 Python 2 还是 Python 3····· 6	2.6.1 PEP 8 编码规范介绍····· 34
第2章 Python生态工具 ····· 9	2.6.2 使用 pycodestyle 检查代码 规范····· 36
2.1 Python 内置小工具····· 9	2.6.3 使用 autopep8 将代码格式化··· 37
2.1.1 1 秒钟启动一个下载服务器··· 10	2.7 Python 工作环境管理····· 39
2.1.2 字符串转换为 JSON····· 10	2.7.1 使用 pyenv 管理不同的 Python 版本····· 39
2.1.3 检查第三方库是否正确安装··· 11	2.7.2 使用 virtualenv 管理不同的 项目····· 42
2.2 pip 高级用法····· 12	2.8 本章总结····· 43
2.2.1 pip 介绍····· 12	第3章 打造命令行工具 ····· 44
2.2.2 pip 常用命令····· 12	3.1 与命令行相关的 Python 语言 特性····· 44
2.2.3 加速 pip 安装的技巧····· 14	
2.3 Python 编辑器····· 15	

3.1.1	使用 <code>sys.argv</code> 获取命令行参数	45	4.2.1	正则表达式语法	86
3.1.2	使用 <code>sys.stdin</code> 和 <code>fileinput</code> 读取标准输入	46	4.2.2	利用 <code>re</code> 库处理正则表达式	87
3.1.3	使用 <code>SystemExit</code> 异常打印错误信息	48	4.2.3	常用的 <code>re</code> 方法	90
3.1.4	使用 <code>getpass</code> 库读取密码	49	4.2.4	案例：获取 HTML 页面中的所有超链接	94
3.2	使用 <code>ConfigParse</code> 解析配置文件	49	4.3	字符集编码	94
3.3	使用 <code>argparse</code> 解析命令行参数	52	4.3.1	编码历史	95
3.3.1	<code>ArgumentParse</code> 解析器	52	4.3.2	UTF-8 编码	96
3.3.2	模仿 MySQL 客户端的命令行参数	54	4.3.3	从字符集的问题说起	98
3.4	使用 <code>logging</code> 记录日志	55	4.3.4	Python 2 和 Python 3 中的 Unicode	99
3.4.1	日志的作用	55	4.4	Jinja2 模板	101
3.4.2	Python 的 <code>logging</code> 模块	56	4.4.1	模板介绍	102
3.4.3	配置日志格式	56	4.4.2	Jinja2 语法入门	102
3.5	与命令行相关的开源项目	59	4.4.3	Jinja2 实战	108
3.5.1	使用 <code>click</code> 解析命令行参数	59	4.4.4	案例：使用 Jinja2 生成 HTML 表格和 XML 配置文件	112
3.5.2	使用 <code>prompt_toolkit</code> 打造交互式命令行工具	61	4.5	本章总结	116
3.6	本章总结	64	第5章 Linux系统管理		117
第4章 文本处理		65	5.1	文件读写	118
4.1	字符串常量	66	5.1.1	Python 内置的 <code>open</code> 函数	118
4.1.1	定义字符串	66	5.1.2	避免文件句柄泄露	119
4.1.2	字符串是不可变的有序集合	68	5.1.3	常见的文件操作函数	120
4.1.3	字符串函数	71	5.1.4	Python 的文件是一个可迭代对象	121
4.1.4	案例：使用 Python 分析 Apache 的访问日志	79	5.1.5	案例：将文件中所有单词的首字母变成大写	122
4.1.5	字符串格式化	83	5.2	文件与文件路径管理	123
4.2	正则表达式	85	5.2.1	使用 <code>os.path</code> 进行路径和文件管理	123