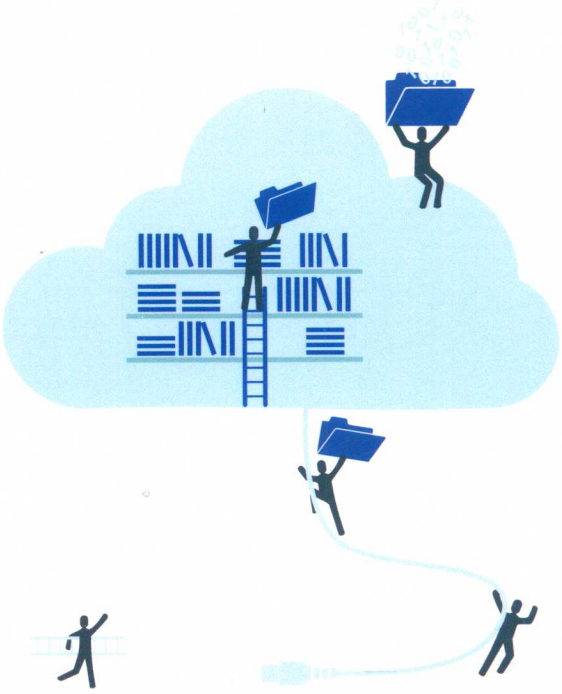
Broadview®
www.broadview.com.cn

Linux开源存储技术大图谱

OpenSDS、OpenStack存储、
容器及k8s存储、OPNFV存储等，全方位、深层次

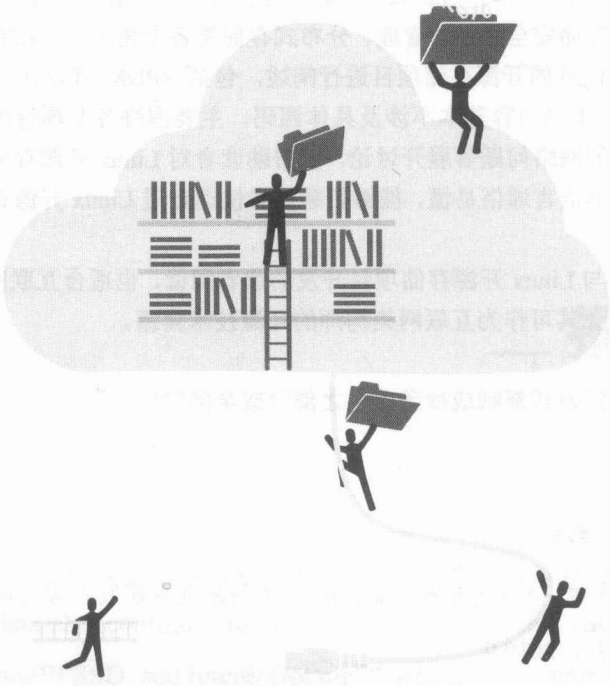


Linux开源存储全栈详解

从Ceph到容器存储

英特尔亚太研发有限公司　编著

中国工信出版集团
電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn



Linux开源存储全栈详解

从Ceph到容器存储

英特尔亚太研发有限公司　编著

電子工業出版社
Publishing House of Electronics Industry
北京•BEIJING

内 容 简 介

本书致力于帮助读者形成有关 Linux 开源存储世界的细致拓扑。本书从存储硬件、Linux 存储栈、存储加速、存储安全、存储管理、分布式存储等各个角度与层次展开讨论，同时对处于主导地位的、较为流行的开源存储项目进行阐述，包括 SPDK、ISA-L、OpenSDS、Ceph、OpenStack、容器等。本书内容基本不涉及具体源码，主要围绕各个项目的起源与发展、实现原理与框架、要解决的网络问题等展开讨论，以帮助读者对 Linux 开源存储技术的实现与发展形成清晰的认识。本书语言通俗易懂，能够带领读者快速走进 Linux 开源存储的世界并做出自己的贡献。

本书适合希望参与 Linux 开源存储项目开发的读者阅读，也适合互联网应用的开发者、架构师和创业者参考，尤其可作为互联网架构师的开源技术典籍。

图书在版编目（CIP）数据

Linux 开源存储全栈详解：从 Ceph 到容器存储 / 英特尔亚太研发有限公司编著.

北京：电子工业出版社，2019.9

ISBN 978-7-121-36979-7

Ⅰ. ①L… Ⅱ. ①英… Ⅲ. ①Linux 操作系统 Ⅳ.①TP316.85

中国版本图书馆 CIP 数据核字（2019）第 127853 号

责任编辑：孙学瑛　　特约编辑：田学清
印　　刷：北京季蜂印刷有限公司
装　　订：北京季蜂印刷有限公司
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：24.75　　字数：476 千字
版　　次：2019 年 9 月第 1 版
印　　次：2019 年 9 月第 1 次印刷
定　　价：99.00 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）88254888，88258888。

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

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。

作者简介

任桥伟：从事Linux内核、OpenStack、Ceph等开源项目的开发，著有《Linux内核修炼之道》《Linux那些事儿》系列。

李晓燕：活跃于Cinder和Ceph项目，具有多年存储领域经验。

程盈心：Ceph社区的活跃贡献者，专注于分布式系统的分析与优化。

马建朋：Ceph社区的活跃贡献者。

刘春梅：哈尔滨工业大学自动控制专业博士，目前在美国硅谷英特尔工作。从事过网络安全、虚拟化、终端安全、云计算等领域的工作。

尚德浩：Ceph社区的活跃贡献者。

胡伟：从事云计算和边缘计算相关工作。在OpenStack、Ceph和Edge Computing领域支持客户技术方案落地，参与业界多项前沿云计算相关技术验证和评估工作。

杨子夜：从事存储软件开发和优化工作。在虚拟化、存储、云安全等领域拥有诸多专利提交，其中21个专利已经被专利局授予（其中14个在美国，7个在中国）。

曹刚：从事存储软件的开发和优化工作，现为英特尔开发经理。

刘长鹏：从事存储软件和虚拟化研发工作。

刘孝冬：从事存储软件研发及存储相关算法优化的工作。

惠春阳：从事存储软件研发及存储相关算法优化的工作。

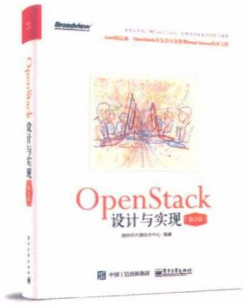
万群：从事测试领域的研究及实践近十年，对测试方法及项目管理有相当丰富的经验。

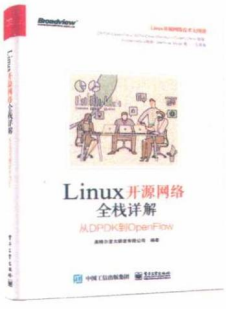
闫亮：从事存储软件的测试和优化工作。

周雁波：在英特尔实习期间，从事存储软件的开发和优化工作。

徐雯昀：在英特尔实习期间，担任存储技术文档工程师，负责SPDK技术文档方面的工作。

好书力荐





拒绝堆砌臃肿，支持纯正原创

投稿邮箱：sxy@phei.com.cn

加入本书读者 QQ 群 465398813，以书会友，资源共享！

推荐序一

With an ever-increasing growth of data in both business and consumer uses, the need for storage of the data also continues to grow at an unprecedented pace. Intel technologies such as Intel® Optane™ SSD, and Intel® Optane™ persistent memory serve as examples of disruptive innovation, providing dramatically improved performance and entire new functionality to the storage industry.

Infrastructure software, storage, computing and network functions will serve a broad range of application platform requirements for 5G, Artificial Intelligence, Autonomous Driving, and Big Data Analytics. Open source storage software is critically important in developing and delivering innovative solutions. This book introduces open source storage technologies and techniques for software-defined storage solutions covering both hardware and software. Intel is a leading contributor of open source software and storage hardware IP and products, including contributions for open source storage software development, such as Intelligent Storage Acceleration Library (ISA-L), Storage Performance Development Kit (SPDK), and the Ceph project.

In China, open source and software-defined storage have been well adopted and applied across many different types of application scenarios. A strong ecosystem of storage software developers and solutions providers have been created and are delivering value to

the market through their innovations. We hope the technologies and techniques covered in this book will be useful to the open source communities in China and the rest of the world.

Sandra Rivera

Senior Vice President and General Manager

Network Platforms Group

Intel 5G Executive Sponsor

推荐序二

The evolution of storage has been dramatic. As computing has become an integral part of our daily activities, storage media has had to evolve to keep pace, moving from tape and mechanical hard disk, to solid state hard disk, NAND, and now emerging 3D XPoint technologies.

Demand is only expected to increase. The data boom created by connected devices coupled with growth areas such as Artificial Intelligence and Machine Learning bring higher expectations — greater capacity, scalability, and faster access, not to mention increased availability and reliability. These challenges have resulted in the rise of distributed storage systems and cloud-based services, to help providers meet those demands. Software plays a key role in this transformation. The emergence of software-defined storage enables more scalable and flexible solutions, as well as greater levels of management and orchestration. Offerings such as virtual machine and container persistence storage help meet a growing range of business needs.

Open source provides the foundation for modern storage systems and a platform for rapid innovation. Intel has been active in the open source software community for 20+ years, and our contributions to storage technologies such as Intel® Storage Acceleration Library (ISA-L), Storage Performance Development Kit (SPDK), and the Ceph project

help accelerate performance for storage virtualization platforms.

As a leader in developing storage hardware and software, Intel recognizes the importance of the Chinese ecosystem to the growth of the storage market. It is with great pride we present this book as a resource to the China storage community.

Imad Sousou

Corporate Vice President, Intel Corporation

General Manager, Intel System Software Products

前　言

自 1991 年 Linux 诞生，时间已经走过了近三十年。即将而立之年的 Linux 早已没有了初生时的稚气，它正在各个领域展示自己成熟的魅力。

以 Linux 为基础，各种开源生态，如网络、存储都出现了。而生态离不开形形色色的开源项目，在人人谈开源的今天，一个又一个知名的开源项目正在全球快速生长。当然，本书的主题仅限于 Linux 开源存储生态，面对其中一个又一个扑面而来且快速更迭的新项目、新名词，我们会有一种紧迫感，想去了解它们背后的故事，也会有一定的动力想要踏上 Linux 开源存储世界的旅程。而无论是否强迫，面对这样的一段旅程，我们心底浮现的最为愉悦的开场白或许应该是："说实话，我学习的热情从来都没有低落过。Just for Fun." 正如 Linus 在自己的自传 *Just for Fun* 中所希望的那样。

面对 Linux 开源存储这么一个庞大而又杂乱的世界，让人最为惴惴不安的问题或许便是：我该如何更快、更好地适应这个全新的世界？人工智能与机器学习领域里研究的一个很重要的问题是，"为什么我们小时候有人牵一匹马告诉我们那是马，于是之后我们看到其他的马就知道那是马了？"针对这个问题的一个结论是：我们在头脑里形成了一种生物关系的拓扑结构，我们所认知的各种生物都会放进这个拓扑结构里，而生物不断成长的过程就是形成并完善各种各样的或树形、或环形等拓扑结构的过程，并以此来认知我们所面对的各种新事物。

由此可见，或许我们认知 Linux 开源存储世界最快也最为自然的方式就是努力在脑海里形成它的拓扑结构，并不断细化，比如这个生态包括什么样的层次，每个层次里又有什么样的项目去实现，各个项目又实现了哪些服务及功能，这些功能又是以什么样的方式实

现的，等等。对于感兴趣的项目，我们可以更为细致地去勾勒其中的脉络，就好似我们头脑里形成的有关一个城市的地图，它有哪些区，区里又有哪些标志性建筑及街道，对于熟悉的地方，我们甚至可以将它的周围放大并细化到一个微不足道的角落。

本书的组织形式

本书正是为帮助读者形成有关 Linux 开源存储世界的细致拓扑而组织的。

第 1 章主要对 Linux 开源存储的生态进行整体描述，包括开源存储领域研究的热点方向、相关的开源基金会等。

第 2 章从存储硬件的角度介绍了存储技术的发展历史，包括存储介质的进化、存储协议的更新等。

第 3 章作为整个 Linux 开源存储世界的基础，描述了 Linux 存储栈（Linux Storage Stack），对 I/O 在 Linux 内核里的处理流程及所涉及的主要模块进行介绍。

第 4 ~ 9 章的内容分别从存储加速、存储安全、存储管理与软件定义存储、分布式存储与 Ceph、OpenStack 存储、容器存储等角度与层次对处于主导地位的、较为流行的项目进行介绍，以帮助读者对相应项目形成比较细致的拓扑。

第 4 章讲解了存储领域的加速技术，包括 FPGA、QAT、NVDIMM 等硬件加速技术，以及 ISA-L、SPDK 等开源的软件加速方案。

第 5 章从可用性、可靠性、数据完整性、访问控制、加密与解密等方面讨论了存储安全问题。

第 6 章介绍存储管理与软件定义存储方面的主要开源项目，包括 OpenSDS、Libvirt 等。

第 7 章讨论分布式存储并详细介绍了目前流行的开源分布式存储项目 Ceph 的设计与实现。

第 8 章与第 9 章分别对 OpenStack 与 Kubernetes 两种主要云平台中的存储支持进行讨论。

感谢

作为英特尔的开源技术中心，参与各个 Linux 开源存储项目的开发与推广是再自然不过的事情了。除了为各个开源项目的完善与稳定贡献更多的思考和代码，我们还希望通过

这本书让更多的人更快地融入 Linux 开源存储世界的大家庭。

如果没有 Sandra Rivera（英特尔高级副总裁兼网络平台事业部总经理）、Imad Sousou（英特尔公司副总裁兼系统软件产品部总经理）、Mark Skarpness（英特尔系统软件产品部副总裁兼数据中心系统软件总经理）、Timmy Labatte（英特尔网络平台事业部副总裁兼软件工程总经理）、练丽萍（英特尔系统软件产品部网络与存储研发总监）、冯晓焰（英特尔系统软件产品部安卓系统工程研发总监）、周林（英特尔网络平台事业部中国区软件开发总监）、梁冰（英特尔系统软件产品部市场总监）、王庆（英特尔系统软件产品部网络与存储研发经理）的支持，这本书不可能完成，谨在此感谢他们在本书编写过程中给予的关怀与帮助。

感谢本书编辑孙学瑛老师，从选题到最后的定稿，在整个过程中，都给予我们无私的帮助和指导。

感谢参与各章内容编写的各位同事，他们是李晓燕、程盈心、马建朋、尚德浩、胡伟、刘春梅、任桥伟、杨子夜、曹刚、刘长鹏、刘孝冬、惠春阳、万群、闫亮、周雁波、徐雯昀。为了本书的顺利完成，他们付出了很多努力。

感谢所有对 Linux 开源存储技术抱有兴趣或从事各个 Linux 开源存储项目工作的人，没有你们提供的源码与大量技术资料，本书便会成为无源之水。

目 录