



经典译丛



ELSEVIER
爱思唯尔

信息网络技术与网络科学

软件定义网络 ——原理、技术与实践

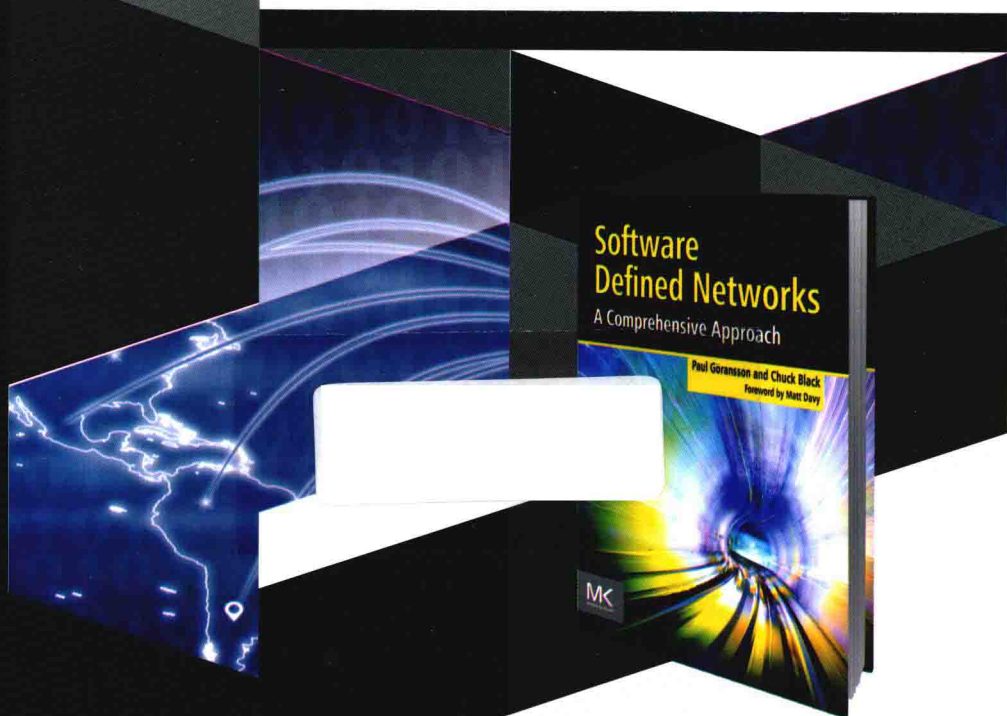
Software Defined Networks

A Comprehensive Approach

【美】Paul Göransson Chuck Black 著

王海 张娟 于卫波 牛大伟 译

Software Defined Networks
A Comprehensive Approach



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

经典译丛·信息网络技术与网络科学

软件定义网络

——原理、技术与实践

Software Defined Networks
A Comprehensive Approach

[美] Paul Göransson Chuck Black 著

王海 张娟 于卫波 牛大伟 译



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

内 容 简 介

软件定义网络(SDN)是由 Emulex 公司提出的一种新型网络创新架构,旨在帮助网络跨入云时代,适应快速变化的业务需求。本书全面介绍了软件定义网络的基本概念、原理和商业应用。全书共 13 章,从多个方面阐释了 SDN,从 SDN 的由来和历史沿革,到 SDN 技术的最新发展情况和未来走向,全面分析了 SDN 和 OpenFlow 的技术原理,分析了 SDN 的开源代码及相关资源,并探讨了 SDN 的应用场景、商业发展及其局限性,是一本视野广阔的 SDN 参考手册。

本书对从事 SDN 技术研发的专业人士、网络运营管理人员、高等院校相关专业的学生以及对 SDN 技术感兴趣的读者都具有一定的参考价值。

Software Defined Networks: A Comprehensive Approach

Paul Göransson, Chuck Black

9780124166752

Copyright © 2014 Elsevier Inc. All rights reserved.

Authorized Simplified Chinese translation edition published by the Proprietor.

Copyright © 2016 by Elsevier (Singapore) Pte Ltd. All rights reserved.

Published in China by Publishing House of Electronics Industry under special arrangement with Elsevier (Singapore) Pte Ltd. This edition is authorized for sale in China Mainland. Unauthorized export of this edition is a violation of Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书简体中文版由 Elsevier (Singapore) Pte Ltd. 授予电子工业出版社在中国大陆发行与销售。未经许可之出口,视为违反著作权法,将受法律之制裁。

本书封底贴有 Elsevier 公司防伪标签,无标签者不得销售。

版权贸易合同登记号 图字: 01-2015-1618

图书在版编目(CIP)数据

软件定义网络:原理、技术与实践/(美)戈朗生(Goransson,P.), (美)布莱克(Black,C.)著;王海等译.

北京:电子工业出版社,2016.3

(经典译丛·信息网络技术与网络科学)

书名原文:Software Defined Networks: A Comprehensive Approach

ISBN 978-7-121-28035-1

I. ①软… II. ①戈… ②布… ③王… III. ①计算机网络-研究 IV. ①TP393

中国版本图书馆 CIP 数据核字(2016)第 007090 号

策划编辑:杨 博

责任编辑:杨 博 特约编辑:张传福

印 刷:三河市双峰印刷装订有限公司

装 订:三河市双峰印刷装订有限公司

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:787×1092 1/16 印张:18 字数:461 千字

版 次:2016 年 3 月第 1 版

印 次:2016 年 3 月第 1 次印刷

定 价:69.00 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 zlts@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

译者序

我们非常高兴地向广大读者推荐本书。

软件定义网络是近年来网络发展的一个重点方向，它与大数据、云计算一起吸引了众多科技工作者、管理者和广大用户的注意力。铺天盖地的宣传与产品已经在市场上涌现，但对于大多数读者来说，仍存在“到底什么是软件定义网络？”“软件定义网络到底好在哪里？”“软件定义网络是否适合我？”的疑问。本书以通俗易懂的语言、全面广泛的图表，精确、全面地解答了上述问题。

今天，在互联网上从来不缺少有关软件定义网络的信息，包括各种形式的博客、播客、微博等。然而，正如读者将在本书里看到的那样，软件定义网络这个名词下面实际上囊括了多种完全不同的技术、体制以及具体的实现系统，大部分针对软件定义网络的介绍都只是讨论了这些技术的一部分，并且由于观察角度和观察对象的不同，难免以偏概全，这也是很多专业人士对软件定义网络的观点大相径庭的原因。市面上关于软件定义网络的书也有不少，但从整体看，有的书侧重描述协议细节，对于技术的整体缺乏阐述，只见树木不见森林；有的又过于宏观，缺乏细节。本书作者 Paul Göransson 博士和 Chuck Black 先生凭借他们多年浸淫在软件定义网络里的工程 and 实践经验里，站在相对公正的立场上巨细靡遗地揭示了软件定义网络的宏大家族，从历史、发展、应用、未来、生态链、周边产业等多个角度介绍了软件定义网络。清晰准确地告诉我们软件定义网络是什么、不是什么，它能帮助我们跳入软件定义网络这一浩瀚的海洋，同时又能避开迷思的漩涡。

本书宏观与微观兼具，树木与森林并览，读此书，您会惊讶于作者从如此多的视角来观察一项技术，给读者提供了充足的实例来认识一项技术的全貌。

任何对网络和软件定义网络感兴趣的读者，都能轻而易举地理解本书的内容，包括本科生和研究生、网络专业人员和 IT 经理。本书内容包容性很强，不需要广泛的网络知识，就能让读者充分了解其价值。本书通过对网络发展史的简介及其技术的描述，再加上对 OpenFlow 和软件定义网络行业全景的描绘，为读者提供了丰富的背景知识。如果您是一位网络架构师或 IT 经理，可能经常会在市面上发现多种解决方案，它们的名字都是软件定义网络，但是其细节却大相径庭，此时您往往会感到困惑，不知道哪种方案最适合您。阅读本书可以为您理解和评估市场上相互竞争的各种软件定义网络技术手段提供坚实的基础。

本书内容涵盖之广，超出您想象。甚至包括软件定义网络对未来 IT 职业的影响。作为一个 IT 从业者，您的职业生涯会受到软件定义网络的影响吗？您应该如何改变自己以适应这个发展大潮？建议您阅读本书。

限于水平，翻译不妥和错误之处在所难免。敬请广大读者批评指正。来信请发至 haiwang@ieee.org。

译者

序 言

什么是软件定义网络？它为什么会风靡网络界？这似乎是一个十分简单的问题，但是却让我花了将近三年的时间去世界各地为上至大型网络运营商的首席技术官，下至本科工程专业的学生反复解答此问题。

当我在2010年7月首次公开谈论SDN的核心协议OpenFlow时，几乎还没有人听说过OpenFlow或软件定义网络，即使在大学校园的网络技术社区里也是如此。然而仅仅过了10个月，在2011年5月，OpenFlow已然成为Interop Las Vegas的热点话题，并且在这个展会上有15个供应商展示了自己的OpenFlow原型。此时，以促进和规范SDN为宗旨的开放网络基金会(ONF)已经成立，而每一本网络技术方面的出版物都在讲述着与这种新技术有关的故事。那可真是一段SDN的美好时光！

在2010年至2011年期间，作为美国印第安纳大学的网络架构师，我和大多数网络运营商一起致力于向大家解释OpenFlow以及SDN能够为大型企业及服务提供商的网络提供的价值。与此同时，我也参加了很多会议，接听了无数电话，并且喝着啤酒与来自思科、Juniper(瞻博网络)、惠普、IBM、戴尔、Brocad(博科)这些行业巨头以及其他许多后起之秀的工程师及管理人员展开热烈讨论。我的团队在早期的试点工作中曾与Nicira和Big Switch Networks之间有过合作，这两家公司是最早，也是最受关注的SDN初创公司。我甚至在早于谷歌首次公开谈论其OpenFlow Network项目的两年之前就有幸面晤来自谷歌的研发人员，并就此事交换了意见。

在这段时间里，我不断向人们解释什么是SDN以及它为什么非同凡响且能惠及成千上万的人，从学生到网络工程师，再到高管层，甚至包括一些大学校长。有些人在听了几分钟的讲解后立刻就“领悟”了，此时我往往还没来得及放映第三张幻灯片；而有些人则要上完整个全天课程才能“明白”。比较普遍的现象是，能够最快理解SDN对行业产生潜在影响的人，恰恰是那些能回忆起因特网之前时代的网络界元老，是实际上构筑了这些网络产品的人以及熟悉网络行业的人。

这就是为什么当Paul和Chuck带着本书的写作计划找到我时，我会感到如此兴奋的原因。Paul和Chuck都拥有丰富的网络从业经验，从业史甚至比他们自己认可的经历还要长。他们通过一代又一代的技术构建了各种各样的网络解决方案。Paul有过两个成功的初创企业，因而熟识网络行业中的各项工作。他的博大精深的工作经验对于解释SDN以及定位SDN在过去三十年的计算机网络界的地位，并预测它对未来几年的网络所具有的潜在影响来说绝对是无价之宝。

也许本书的灵感来自Chuck与我的一次谈话，当时Chuck刚编写完他的第一个基于SDN控制器的网络接入控制(NAC)解决方案。Chuck在惠普担任了多年的网络架构师，负责该公司的NAC解决方案，并且对基于“遗留”的技术来构建NAC解决方案的困难之

处有着亲身体验，这些技术包括简单网络管理协议(SNMP)和远程认证拨号用户服务(RADIUS)。当他在数天之内，而不是花了数月的时间，在开源 SDN 控制器上实现了可以正常工作的 NAC 方案后，Chuck 就对 SDN 坚信不疑了。

今天，在因特网上从来不缺有关 SDN 的各种信息，包括博客、播客、微博，等等。甚至还有一本关于 OpenFlow 协议的自助出版读物。然而，对于那些在过去几年未能直接参与 SDN 而又想要迎头赶上的公司来说，在本书出现之前却很难找到一本能够全面详尽地介绍与 SDN 相关信息的书籍。

任何人，只要对网络和 SDN 感兴趣，都不难理解本书的内容，包括本科生、研究生、网络专业人员和 IT 经理。本书内容的自包容性很强，不需要广泛的网络知识就能让读者充分了解其价值。本书通过对网络发展史的概述以及对各种网络技术的介绍，再加上对引发 OpenFlow 和 SDN 出现的网络行业格局的描绘，为读者提供丰富的背景知识。如果你是一位网络架构师或 IT 经理，正在试图比较多个都声称自己是基于 SDN 的，但又似乎是完全不同的解决方案时，本书可以为你理解和评估市场上相互竞争的各种 SDN 方案提供坚实的基础。

在过去三年中，我花了超出自己意愿的时间来回答这个简单的问题，“什么是 SDN？它的价值在哪里？”经常有人通过电子邮件或者在各种会议中告诉我说他们想要更多了解 SDN，并咨询我应该阅读哪些资料。迄今为止，我不得不指点给他们一长串的播客、白皮书、博客以及会议视频，而这些资料对 SDN 的看法非常零散、不连贯，并且往往带有一些偏见。本书的出现，带来了与 SDN 相关的内容全面、详细地介绍，令我深感满意，在此很愿意将此书推荐给大家。

Matt Davy
Tallac Networks 公司
2013 年 10 月 7 日

Matt Davy 世界知名软件定义网络技术专家，在美国印第安那大学担任 InCNTRE、SDN 互操作实验室、网络研究部以及实习和训练部门的执行主任。Matt Davy 是拥有超过十二万用户、十万以太网端口和五千个无线接入点的企业网络的首席设计师，在大型服务提供商网络和企业网络的设计和运行方面拥有 16 年的丰富经验。

前 言

当初我们着手构思这样一本书时，部分动机是出于很难找到一本完整且详细介绍 SDN 的出色读物。虽然身为作者的我们是 SDN 技术的内行，但也不得不承认无法从某个单一来源获得完整的 SDN 信息。我们意识到，对于很多没有直接接触 SDN 但却需要对其有所了解的专业人员而言，这真是一个大问题。因此，泛泛而言，本书的目标就是描述产生 SDN 的环境，勾画 SDN 区别于其他竞争技术的特点，并解释这种新兴技术已展现出的许多重要的商业影响。

为这样一种快速发展的技术撰写其早期著作所面临的一个挑战就是目标的不稳定性。在此期间，SDN 技术进展速度之快使得我们只能争取对其当前状态做出尽量准确、详实的临摹。在本书出版之后，SDN 技术肯定还会继续发展。我们选择了使用“全面的解决方案”作为本书标题的一部分。目前有许多设计思想正处于相互竞争之中，而其创造者都希望赶上 SDN 的潮流。不论读者需要应对的是哪种类型的 SDN 技术或 SDN 技术的哪些方面，至少我们希望他或她通过阅读本书能够在更广泛的 SDN 背景中找到自己的位置。为此，我们试图全面介绍 SDN 的各种定义。在本书中，我们对 SDN 定义的应用没有教条可寻，希望读者能够理解。

有兴趣了解软件定义网络或者对下面任何主题感兴趣的读者都将在阅读本书后有所收获：

- 组网技术
- 交换技术
- 软件定义网络
- OpenFlow
- OpenStack
- 网络虚拟化

软件定义网络是一个正在迅速扩张的广阔领域。虽然我们试图尽可能做到全面和翔实，但感兴趣的读者可能还是需要根据本书提供的参考资料就某些技术专题做更深入的探索。我们假设读者除了对计算机的概念有基本了解之外，并不具备专业知识。计算机编程和计算机组网的某些经验对理解本书的内容会有帮助。本书包含了大量的数据和图表来解释和说明相关的概念。这些图表能够让读者通读全书而不会让他们觉得必须要借助其他参考文献。

建议和纠错

虽然我们已尽可能地小心以减少错误,但不可避免地还是会出现一些失误,并且也可能忽略了某些读者特别感兴趣的话题。我们期待在未来的版本中能够纠正所有的出版错误,并容纳尽可能多的建议。请通过电子邮件发送评论给我们。

chuck_black@tallac.com

作者简介

Paul Göransson 博士是一位有多次创业经历的企业家。他白手起家创建的两家公司都已成功被行业巨头收购:惠普收购了 Qosnetics (1999);思科收购了 Meetinghouse (2006)。Paul 曾任职于安捷伦科技的先进网络部门和思科的无线网络商业部门的高级管理层。作为 Elbrys Networks 公司的创始人兼董事会主席,保罗目前领导该公司的发展战略并负责 Elbrys 的知识产权目录(intellectual property portfolio)。Paul 拥有美国布兰迪斯大学心理学学士学位、美国波士顿大学计算机工程硕士学位以及美国新罕布什尔大学计算机科学博士学位。Paul 还是狂热的马拉松运动员、登山家、铁人三项选手,并且一直是活跃的潜水员和户外爱好者,他曾参加了六次铁人三项比赛和多次超长距离马拉松比赛,并且是美国全国潜水教练员协会(National Association of Underwater Instructors, NAUI)的潜水长。Paul 在法国、阿尔及利亚、委内瑞拉、墨西哥和瑞典等地居住、研究以及工作。Paul 与人合著了“Roaming Securely in 802.11 Networks”(在 802.11 网络里安全漫游)并撰写了许多与计算机网络技术相关的期刊文章,他也经常应邀在各种学术会议上发表演讲。而在业余时间里,Paul 拥有并管理着缅因州南部的一个 130 英亩的肉牛和草料农场。

Chuck Black 在计算机网络领域拥有超过 31 年的工作经验,其中的大部分时间都是在惠普研究与开发实验室工作,此后他成为 Tallac Networks 公司的创始人之一,该公司是一家 SDN 初创公司。Chuck 最近一直在为惠普的网络接入控制和安全领域开发各种网络产品。在这项工作之前,他曾为惠普的软件部门开发网络管理领域的产品。在局域网发展的早期,Chuck 曾是该行业最早的网络拓扑发现算法的作者。Chuck 持有美国加州州立理工大学的计算机科学学士和硕士学位。

致谢

非常感谢我们的家庭这些年来给予的大力支持,特别是在我们撰写这本书时。

如果没有 Tallac Networks 公司的合伙创始人 Bill Johnson、Matt Davy 和 Paul Congdon 博士多年来始终如一的支持,就不可能有本书的问世。他们对 SDN 技术的深刻理解,以及克服各种困难对本书手稿所做的复审工作,再加上他们对本书的其他许多直接贡献,所有这一切都是无价之宝。

我们还要感谢来自荷兰 Sicse 的互联网通信与服务工程师 Niels Sluijs 博士,他对本书的手稿给出了许多令人耳目一新且非常中肯的观点和评论。

我们要感谢远在泰国清迈大学的 Nopadon Juneam, 他帮助我们对书中多幅插图进行最终版本定型。

特别感谢 Elsevier 的 Steve Elliot 和 Kaitlin Herbert 对这个项目的鼓励和支持。

我们还要感谢 Tallac Networks 公司的 Ali Ezzet, 他仔细阅读了书稿的各个章节, 并运用自己的技术专长发现了不少错误, 大大提高了书稿的质量。我们同样赞赏 Elbrys Networks 公司的 Anthony Delli Colli, 是他在商业和市场营销方面为本书提供了许多重要的信息和建议。这些对于本书来说都是不可或缺的贡献。

特别感谢 Helen Göransson 为本书手稿所做的多次精心审阅, 她的努力使本书更具可读性。

我们也非常感谢 Tallac Networks 公司为全书所使用插图的选定所做出的贡献。

Paul Göransson

Chuck Black

目 录

第1章 绪言	1	2.2 成本	24
1.1 分组交换的基本术语	2	2.2.1 开发成本不断增加	25
1.2 历史背景	4	2.2.2 封闭的环境鼓励供应商的市场 圈定	25
1.3 现代数据中心	5	2.2.3 变革的复杂性与阻力	26
1.4 传统交换机体系结构	7	2.2.4 网络运营成本的上升	26
1.4.1 数据、控制和管理平面	7	2.3 SDN 意味着不断的探索和 革新	26
1.4.2 基于软件的路由选择和 桥接	8	2.3.1 现状对目前主流供应商 有利	27
1.4.3 转发表的硬件查找	9	2.3.2 SDN 对科研创新的促进 作用	27
1.4.4 通用可编程的转发规则	9	2.4 数据中心的创新	28
1.5 自治和动态转发表	10	2.4.1 计算和存储虚拟化	28
1.5.1 二层控制	11	2.4.2 当前网络的不足之处	29
1.5.2 三层控制	12	2.5 数据中心的要求	30
1.5.3 是协议大杂烩还是一碗 交换机粥	13	2.5.1 自动化	30
1.6 我们能提高分组转发的 IQ 值吗	16	2.5.2 可扩展性	30
1.7 开源和技术转变	17	2.5.3 多路径	31
1.8 本书的组织结构	18	2.5.4 多租户	31
参考文献	18	2.5.5 网络虚拟化	31
第2章 为什么是 SDN	20	2.6 小结	32
2.1 交换机及控制平面的发展 历程	20	参考文献	32
2.1.1 用软件实现的简单的分组转发和 路由选择	21	第3章 SDN 的起源	34
2.1.2 早期设备的独立性和 自治性	21	3.1 网络技术的发展历程	34
2.1.3 从软件走向硬件	23	3.1.1 大型机网络: 远程终端	34
2.1.4 硬件转发以及用软件实现的 控制	23	3.1.2 点对点对等连接	35
2.1.5 简化设备的需求越来越 强烈	23	3.1.3 局域网	35
2.1.6 控制从设备上的移除	24	3.1.4 桥接网络	35
		3.1.5 路由选择网络	36
		3.2 SDN 的前身	37
		3.2.1 早期的尝试	37
		3.2.2 网络接入控制	38

3.2.3	编排	39	4.4.3	目前 SDN 控制器的实现 ...	63
3.2.4	虚拟管理器网络插件	40	4.4.4	SDN 控制器的潜在问题 ...	64
3.2.5	ForCES: 转发和控制平面的 分离	41	4.5	SDN 应用	65
3.2.6	4D: 集中的网络控制	42	4.5.1	SDN 应用的职责	65
3.2.7	Ethane: 基于控制器的网络 策略	44	4.6	替代的 SDN 方法	66
3.3	软件定义网络的诞生	45	4.6.1	利用现有 API 的 SDN	66
3.3.1	OpenFlow 的出现	45	4.6.2	利用基于虚拟机管理程序的 覆盖网络的 SDN	68
3.3.2	开放网络基金会	46	4.7	小结	70
3.4	维护 SDN 的互操作性	47	参考文献	71	
3.5	开放源代码的积极作用	48	第 5 章	OpenFlow 规范	72
3.5.1	集体的力量	48	5.1	本章使用的术语	73
3.5.2	集体的危险性	49	5.2	OpenFlow 概述	73
3.5.3	SDN 开源代码	49	5.2.1	OpenFlow 交换机	74
3.6	传统机器向 SDN 的演化	50	5.2.2	OpenFlow 控制器	75
3.7	网络虚拟化	50	5.2.3	OpenFlow 协议	76
3.8	我可以把自己的网络称为 SDN 吗	51	5.2.4	控制器与交换机之间的安全 通道	76
3.9	小结	51	5.3	OpenFlow 1.0 和 OpenFlow 的 基本概念	77
参考文献	51		5.3.1	端口和端口队列	78
第 4 章	SDN 的工作原理	53	5.3.2	流表	78
4.1	SDN 的基本特点	53	5.3.3	分组匹配	79
4.1.1	平面分离	53	5.3.4	动作与分组转发	80
4.1.2	简化的设备和集中的 控制器	54	5.3.5	控制器和交换机之间的 消息	82
4.1.3	网络自动化和虚拟化	54	5.3.6	举例: 控制器对流表的 编程	84
4.1.4	开放性	55	5.3.7	举例: 基本的分组转发	86
4.2	SDN 的工作原理	55	5.3.8	举例: 交换机向控制器转发 分组	86
4.3	SDN 网络设备	57	5.4	OpenFlow 1.1 新增功能	88
4.3.1	流表	58	5.4.1	多级流表	88
4.3.2	软件 SDN 交换机	59	5.4.2	群组	90
4.3.3	硬件 SDN 设备	59	5.4.3	MPLS 和 VLAN 标签的 支持	91
4.3.4	SDN 设备的现有实现	61	5.4.4	虚拟端口	91
4.3.5	流的数量调整	61	5.4.5	控制器连接失效	92
4.4	SDN 控制器	61			
4.4.1	SDN 控制器的核心模块 ...	62			
4.4.2	SDN 控制器接口	63			

5.4.6	举例：使用多级流表的转发	92	6.2.2	适用于 SDN-via-API 的 API	117
5.4.7	举例：使用 V.1.1 的群组多播	94	6.2.3	REST：一种不同的 API 协议	117
5.5	OpenFlow 1.2 的新增功能	94	6.2.4	SDN-via-API 举例	118
5.5.1	支持可扩展匹配	95	6.2.5	SDN-via-API 的评估	119
5.5.2	支持可扩展的 SET_FIELD 分组重写	95	6.3	基于虚拟机管理程序的覆盖网络 SDN	120
5.5.3	PACKET_IN 中可扩展的上下文表达式	95	6.3.1	覆盖网络中的控制器	120
5.5.4	多控制器	96	6.3.2	覆盖网络的操作过程	121
5.5.5	举例：经过 SP 网络的桥接 VLAN	96	6.3.3	SDN-via-overlay 举例	122
5.6	OpenFlow 1.3 的新增功能	97	6.3.4	SDN-via-overlay 的评估	123
5.6.1	功能协商的重构	98	6.4	利用开放设备的 SDN	124
5.6.2	支持更灵活的表失配	98	6.5	网络功能虚拟化	125
5.6.3	每流计量	99	6.6	各种替代方法之间的重叠与比较	126
5.6.4	每连接的事件过滤	100	6.7	小结	127
5.6.5	辅助连接	100		参考文献	128
5.6.6	PACKET_IN 消息中的 Cookie	101	第 7 章	数据中心的 SDN	129
5.6.7	运营商骨干网桥接的标签处理	102	7.1	数据中心的定义	129
5.6.8	举例：利用计量段强制实施的 QoS	102	7.2	数据中心的需求	131
5.7	OpenFlow 的局限性	103	7.2.1	克服当前网络的限制	131
5.8	小结	103	7.2.2	增加、移动和删除资源	134
	参考文献	104	7.2.3	故障恢复	134
第 6 章	SDN 的替代定义	105	7.2.4	多租户	134
6.1	OpenFlow 潜在的缺点	105	7.2.5	流量工程和路径效率	135
6.1.1	变化太大、太快	106	7.3	数据中心的隧道技术	135
6.1.2	单点故障	107	7.3.1	虚拟扩展局域网	136
6.1.3	性能和规模	110	7.3.2	使用通用路由封装的网络虚拟化	137
6.1.4	深度分组检测	112	7.3.3	无状态传输隧道	138
6.1.5	状态流意识	114	7.4	数据中心的路径技术	138
6.1.6	结论	115	7.4.1	通用多路径路由选择	139
6.2	SDN-via-API	115	7.4.2	多生成树协议	139
6.2.1	网络设备中的传统 API	116	7.4.3	最短路径桥接	139
			7.4.4	等价多路径	140
			7.4.5	SDN 和最短路径的复杂性	141
			7.5	数据中心的以太网矩阵	141

7.6	数据中心的 SDN 应用场景	142	8.6.2	SDN 应用于服务器负载 均衡	169
7.6.1	克服当前网络的限制	143	8.6.3	SDN 用于防火墙	170
7.6.2	增加、移动和改变网络 资源	144	8.6.4	SDN 应用于入侵检测	171
7.6.3	故障恢复	145	8.7	光传输网络	172
7.6.4	多租户	146	8.7.1	SDN 应用于光网络	172
7.6.5	流量工程和路径有效性	146	8.8	SDN 与 P2P/覆盖网络的 比较	174
7.7	比较开放 SDN 与 SDN-via-overlay 在数据中心的表现	148	8.9	小结	174
7.7.1	SDN-via-overlay	148	参考文献	175	
7.7.2	开放 SDN	148	第 9 章	SDN 生态圈合作伙伴	176
7.8	现实世界的数据中心实现	149	9.1	学术研究机构	177
7.9	小结	149	9.1.1	学术界对 SDN 的主要 贡献	178
参考文献	150		9.2	产业研究实验室	178
第 8 章	非数据中心环境下的 SDN	151	9.3	网络设备制造商	179
8.1	广域网	153	9.4	软件供应商	181
8.1.1	SDN 在广域网中的应用	154	9.5	白盒交换机	183
8.1.2	举例: 谷歌广域网的 MPLS LSP	155	9.6	商用芯片生产商	184
8.2	服务提供商和电信运营商 网络	156	9.7	原始设备制造商	185
8.2.1	SDN 在服务提供商和电信 运营商网络中的应用	158	9.8	企业	185
8.2.2	举例: MPLS-TE 和 MPLS VPN	159	9.9	标准机构和产业联盟	186
8.2.3	举例: 服务提供商的 云爆发	159	9.9.1	开放网络基金	187
8.3	园区网	160	9.9.2	OpenDaylight	187
8.3.1	园区网中的 SDN: 策略的 应用	161	9.9.3	OpenStack	188
8.3.2	园区网中的 SDN: 设备与 用户安全	161	9.9.4	IETF 和 IEEE	189
8.3.3	园区网中的 SDN: 流量 抑制	163	9.10	小结	189
8.4	酒店网	165	参考文献	190	
8.5	移动网络	165	第 10 章	SDN 的应用	192
8.5.1	SDN 应用于移动网络	166	10.1	前期工作	192
8.6	在线网络功能	168	10.2	反应式应用和先应式应用的 比较	193
8.6.1	NFV 与 SDN 的比较	168	10.2.1	术语	194
			10.2.2	反应式 SDN 应用	194
			10.2.3	先应式 SDN 应用	196
			10.3	一些简单 SDN 应用的 分析	198

10.4	一个简单的 Java 反应式应用	198	11.11	举例: SDN 开源代码的应用	233
10.4.1	主机名黑名单	199	11.12	小结	235
10.4.2	IP 地址的黑名单处理	199		参考文献	236
10.4.3	黑名单: 监听器	201	第 12 章	商业影响	237
10.4.4	黑名单: 分组处理器	202	12.1	一切皆服务	237
10.4.5	黑名单: 流管理器	204	12.2	市场规模	238
10.5	控制器的背景	207	12.3	SDN 厂商分类	238
10.6	使用 Floodlight 控制器	207	12.3.1	服务器虚拟化传统运营商和 SDN	239
10.6.1	Java API	208	12.3.2	增值经销商	239
10.6.2	RESTful API	208	12.4	对传统网络设备制造商的影响	240
10.7	使用 OpenDaylight 控制器	209	12.4.1	保护市场份额	240
10.7.1	Java API	209	12.4.2	适度创新	240
10.7.2	RESTful API	209	12.5	对企业客户的影响	241
10.8	使用思科的 XNC 控制器	210	12.5.1	降低设备成本	242
10.8.1	RESTful API	210	12.5.2	避免混乱	242
10.9	使用惠普控制器	210	12.6	在网络行业中引发的风暴	243
10.9.1	Java API	210	12.6.1	大型网络设备制造商变得越来越少	243
10.9.2	RESTful API	211	12.6.2	迁移到云计算	243
10.10	关于交换机的考虑	211	12.6.3	渠道变化	244
10.11	创建网络虚拟隧道	212	12.7	风险投资	244
10.12	数据中心的流卸载	213	12.8	重要的 SDN 收购案	245
10.13	用于园区网的接入控制	215	12.8.1	VMware	246
10.14	服务提供商的流量工程	216	12.8.2	Juniper	246
10.15	小结	217	12.8.3	Brocade	247
	参考文献	218	12.8.4	思科	247
第 11 章	SDN 开源代码	219	12.9	SDN 初创公司	248
11.1	本章特定术语	220	12.9.1	OpenFlow 中坚力量	249
11.2	开源许可证相关事宜	221	12.9.2	非 OpenFlow 的白盒企业	250
11.3	SDN 开源代码用户的特征	223	12.9.3	OpenFlow ASIC?	250
11.4	OpenFlow 的源代码	224	12.9.4	数据中心网络虚拟化	251
11.5	交换机实现	225	12.9.5	WAN 网络虚拟化	252
11.6	控制器实现	226	12.9.6	网络功能的虚拟化	252
11.7	SDN 应用	229	12.9.7	光交换	252
11.8	编排和网络虚拟化	230			
11.9	仿真、测试和工具	231			
11.10	OpenStack	231			

12.9.8 网络边缘的移动功能 和 SDN	253	13.2.2 对网络应用编程技术 ...	261
12.10 职业影响	253	13.2.3 安全应用	262
12.11 小结	254	13.2.4 移动网络中的漫游	263
参考文献	255	13.2.5 移动网络中的流量工程 ...	264
第 13 章 SDN 的未来	257	13.2.6 节能	265
13.1 现状	257	13.2.7 支持 SDN 的交换机 芯片	266
13.2 开放 SDN 潜在的创新 应用	259	13.3 小结	267
13.2.1 管理非传统的物理层 链路	260	参考文献	268
		附录 缩略语	270

Chapter

1

绪 言

科技文章的作者很少能在最近一期主流新闻杂志的头条新闻中看到自己所要撰写的主题，而围绕软件定义网络(Software Defined Networks, SDN)刮起的强劲风暴却使该主题登上了头条新闻^[1]。现代计算机网络已经演变成一个对管理极具挑战性的复杂怪兽，并且仍然在不断膨胀以满足现今环境下的诸多要求。SDN 代表了试图解决当前网络范式存在问题的一种新方法。SDN 是将软件交换机应用于现代数据网络的创新方式。SDN 向着高度可扩展的集中式网络控制体系结构发展的举动更适合于当前百万量级数据中心里普遍存在的超大规模网络。SDN 从一开始就被设计为执行细粒度的流量转发判决，而不是把特定于应用程序的转发行为硬塞进并不适合这项任务的传统体系结构之中。人们对 SDN 的兴趣已远远超越了学术研究及工程技术社区对这种崭新的网络交换技术的关心程度。如果 SDN 的技术承诺得以实现，那么它所代表的不啻于网络行业的结构性转变，老牌的行业巨头很有可能会被拉下台，而被转嫁给消费者的成本也有可能暴跌。不过，人们对此的期待也确实有一定程度的夸张，我们不仅需要了解这种新的网络模型所具有的潜能，同时也应当了解其局限性，这一点很重要。本书尽量从技术的角度解释 SDN 是如何工作的，并综述哪些网络应用是它所适合的，哪些并不适合，我们还要教大家如何在此技术之上构建定制的应用，并从网络业务本身来探讨 SDN 具有的影响力。

作为绪言，本章介绍了与当前互联网交换机的基本概念相关的背景知识，包括对数据平面、控制平面和管理平面的定义和讨论。这些概念是理解 SDN 如何通过与传统交换机体系结构截然不同的方式实现其核心功能的关键。本章还描述了目前的实现是如何进行转发判决的，并说明这种缺乏灵活性的转发判决使得网络管理员很难根据不同条件对网络进行优化。本章继而通过举例说明为什么更加灵活的转发判决可以大大提高现有交换机的业务多样性，并解释为什么将控制平面从交换机中独立出来，使其成为一个独立、开放平台的控制器之后，就能够提高转发判决的灵活性。通过将其与 Linux 操作系统相类比之后得出的结论是：Linux 操作系统因充分利用了开源的开发社区而得到飞速发展，倘若将相同的模式应用到互联网交换机的控制平面开发上，也会带来同样的效果。



接下来先介绍一些分组交换技术的基本术语，这些术语将贯穿全书，之后对分组交换技术及其发展历程进行概述。

1.1 分组交换的基本术语

本节定义了与分组交换技术相关的一些基本术语，这些术语的使用将贯穿全书。我们的惯例是用斜体字来强调新术语的首次使用。对于本节未定义的那些专业性更强的术语，将会在它们首次使用时再进行定义。许多分组交换术语和短语对不同群体而言有不同的含义，本书尝试使用这些术语和短语普遍被人们所接受的定义。缩略语也会在首次使用时定义并强调，本书附录中按字母顺序列出了书中所有用到的缩略语。高级别读者可以跳过本节，而其他读者可能希望先通读本书的全部内容，然后再回头参考特定的术语。

在解释 SDN 与传统分组交换网络有何不同时，术语是一个很重要的参考框架，虽说 SDN 在某种程度上已经背离了这些传统的概念或者从根本上改变了其含义。在阅读本书的过程中，我们鼓励读者随时回顾对这些术语的定义，并思考何时该术语的含义在 SDN 中保持不变，何时 SDN 需要更细致的定义，以及何时对 SDN 的讨论需要用到全新的词汇。

广域网(Wide Area Network, WAN)是指覆盖广阔地理区域范围的网络，通常会跨越多个城区。

局域网(Local Area Network, LAN)是指覆盖有限地理区域范围的网络，面积一般不会超过几千平方米。

城域网(Metropolitan Area Network, MAN)是填补局域网和广域网之间空隙的网络。人们之所以开始使用这个术语，是因为局域网和广域网最初的区别不仅仅体现在地理覆盖范围上，它们所使用的传输技术及速度也不同。随着既有类似局域网的速度和接入控制机制，又能够服务于城市大部分区域的技术的出现，人们开始使用术语“城域网”来表示这样一种既不同于大型局域网，也不同于小型广域网的新实体。

无线局域网(Wireless Local Area Network, WLAN)是以空气为传输媒体的局域网。在无线网络中，任意两个设备之间的最大距离通常在 50 米左右。虽然无线通信也可以不使用空气作为传输媒体，但在本书中我们不考虑这种情况。

物理层(Physical Layer)是计算机网络开放系统互连(Open Systems Interconnection, OSI)^[10]七层模型的最低层，由基本的硬件传输技术组成，这些技术被用来在网络中传输数据比特。

数据链路层(Data Link Layer)是 OSI 模型的倒数第二层。这一层所提供的是同一个网段内数据从一个设备传送到另一个设备的能力。为了清楚起见，我们在此把一个局域网段等同于一个冲突域。局域网段的严格定义是网络设备之间的电气或光学连接。在我们对数据链路层的定义中，通过中继器链接的多个网段也被当作是一个局域网段。网段的例子包括像以太网这样的单个局域网，或者是广域网中相邻节点之间的点对点的通信