

重构网络

SDN架构与实现

杨泽卫 李呈 著

2008
OpenFlow

2009
SDN

2011
ONF

2012
Nicira&B4

2013
OpenDaylight

2014
White Box Switches

2015
P4

2016
SD-WAN

重构网络

SDN架构与实现

杨泽卫 李呈 著

电子工业出版社

Publishing House of Electronics Industry

北京•BEIJING

内 容 简 介

身处 SDN 发展浪潮, 笔者真切地感受到了这场技术革命的到来。为了自我总结, 也为了帮助他人, 因此决定写这本书。

本书首先介绍了 SDN 的定义、SDN 出现的原因、SDN 发展的历史和标志性事件、SDN 南向协议、SDN 控制平面和数据平面等理论知识。然后进一步介绍了如何从零开始搭建 SDN 实验环境及 SDN 与网络虚拟化的结合等 SDN 应用案例内容。最后梳理了 SDN 对学术界和工业界产生的影响, 探讨了我们应该如何应对这场 SDN 变革。此外, 附录中两篇拟人化的故事还从感性的角度对 SDN 进行了介绍。

本书适合 SDN 初学者和进阶者。希望本书的内容能够给读者带来一些帮助, 成为 SDN 学习者相互讨论、学习的舞台。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有, 侵权必究。

图书在版编目(CIP)数据

重构网络: SDN 架构与实现 / 杨泽卫, 李呈著. —北京: 电子工业出版社, 2017.4
ISBN 978-7-121-31042-3

I. ①重… II. ①杨… ②李… III. ①计算机网络—网络结构—研究 IV. ①TP393.02

中国版本图书馆 CIP 数据核字 (2017) 第 044613 号

策划编辑: 张春雨

责任编辑: 徐津平

印 刷: 三河市良远印务有限公司

装 订: 三河市良远印务有限公司

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: 16.75 字数: 307 千字

版 次: 2017 年 4 月第 1 版

印 次: 2017 年 4 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, 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

推荐序

如今，距离我的那本《深度解析 SDN》一书出版已经过去了三年时间，在这短短的三年中，SDN 领域发生了翻天覆地的变化。一大批 Startup 公司在这个领域崭露头角，甚至久已没有新鲜血液的交换芯片领域也出现了新的身影。各个传统巨头也纷纷推出了形态各异的 SDN 产品，不仅仅是网络设备商，还包括一些传统的 IT 厂商。更重要的是，以网络虚拟化、SDWAN 等为代表的典型 SDN 应用纷纷落地。SDN 时代已经到来！

但是对于很多人来说，很多问题仍然看不清楚，到底什么是 SDN？市场上那么多 SDN 产品，哪些是真的，哪些是假的？SDN 到底能够解决哪些传统网络解决不了的问题？SDN 能够全面取代传统网络吗？传统网络应该如何向 SDN 网络逐步迁移？哪些场景是适合 SDN 的典型场景？这些是每一个网络从业者和最终用户都非常关心的市场层面的问题。而在技术层面，同样有很多富有争议的问题。OpenDaylight 或者 ONOS，甚至是其他一个什么控制器，谁会胜出？OpenFlow 前途命运如何？标准南向接口是否是 SDN 应该追求的方向？控制是否真的应该完全从转发面分离？可编程的 P4 是网络未来的方向吗？是否应该有及是否会有真正的 SDN 交换芯片出现？

我跟本书作者杨泽卫和李呈都有数面之缘并都有过交流，也看过不少他们两个人写的文章，两个人都从不同的层面上对 SDN 做了不少卓有成效的研究和实践工作。他们能把工作学习过程中的所见所思所做总结出来，帮助读者来寻找上述问题的答案，不得不说是一件幸事。读者是否能找到所有的答案并不重要，甚至他们的观点是否全部都正

确也并不重要，重要的是，读者可以从他们的工作中系统地学习 SDN 相关知识，并从中得到启发，引出更多的思考，这就是他们这本书最大的意义。

张卫峰

盛科网络 SDN 白牌交换机 CTO

前言

SDN(Software-Defined Networking, 软件定义网络)起源于斯坦福大学 Nick McKeown 教授的 Clean Slate 项目,其目标是重新定义网络体系结构(Reinvent the Internet),诞生至今已经快十年。作为一种新的网络体系结构,SDN 已经掀起了一场网络变革的技术浪潮,对网络学术界和工业界的发展都产生了巨大的冲击:OpenFlow 的论文至今已经被引用 4951 次;开源 SDN 控制器平台 OpenDaylight 已经发布了 5 个版本,拥有超过 600 多名开发者,完成了超过 30000 多次代码更新;开源社区 OSS(Open Source SDN)也已经发布了 20 多个来自开发者的开源 SDN 项目;传统网络设备厂商和运营商都在面向 SDN 重建自己的产品体系,大多数 SDN 初创公司都在各自领域交付成熟的产品方案。而作为学习者的我们,也需要做好准备,迎接这场技术变革。

但是学习和研究 SDN 绝非易事。虽然我们每天都能接触到大量的 SDN 学习资料,但对于初入 SDN 领域的学生和工程师而言,仍然会不知所措。比如,关于 SDN 定义的描述就有多种解读:开源组织 ONF(开放网络基金会)的 Open SDN,业界厂商的 Vendor SDN,甚至还有 SDx(Software Defined Everything)的概念。这些概念哪一种才是正确的?此外,SDN 发展至今诞生了众多不同类型的开源项目:SDN 网络模拟器、SDN 软件交换机、网络虚拟化平台、SDN 控制器测试工具和 OpenFlow 交换机测试工具等。如何去学习它们?如果有一本书能提供清晰的 SDN 学习路线:梳理 SDN 的定义、SDN 发展历程、SDN 关键技术和典型应用案例,同时又介绍如何从零开始实践 SDN,这将会对 SDN 初学者入门和进阶者学习 SDN 提供有效的帮助。

带着这种想法，我们在一年多之前开始编写这本书。本以为依靠学习 SDN 过程中积累的知识和书写博客的经历，就可以轻而易举地完成本书，但事实却大相径庭。从 2015 年 7 月确定本书目录开始，我们每周至少进行一次沟通，直至 2016 年 5 月才陆续完成本书初稿。后续又花费了大量的时间修改初稿，有些章节甚至几经易稿，修改超过二十多次。面对 SDN 这样一个新的技术领域，我们非常谨慎地去选择合适的内容。在写作上，我们努力引用原始技术资料，并在此基础上进行剖析，得出观点和结论。我们希望本书不但能帮助入门者梳理 SDN 领域的知识体系，而且能帮助进阶者挖掘 SDN 领域更深层的信息，在成为 SDN 专业人士之路上贡献我们的一臂之力。

本书包括 8 章内容和 2 个小故事。第 1 章从 SDN 最初的定义出发，讨论了 SDN 出现的原因，详细介绍了 SDN 的发展历史和标志性事件。第 2 章详细介绍了现有的 SDN 南向协议，它是决定 SDN 架构可编程能力的关键，也是当下 SDN 厂商方案竞争的一个焦点。第 3 章从发展历程、系统架构和使用方式 3 个方面详细介绍了 5 个典型的开源 SDN 控制器。第 4 章从 OpenFlow 数据平面开始，深入讨论了现有的 SDN 数据平面模型，并引出一个非常重要的 SDN 数据平面概念“通用可编程数据平面”。第 5 章介绍了从零开始实现 SDN 所需要掌握的诸多开源工具，帮助读者快速入门 SDN 实践。第 6 章讨论了现有的 SDN 应用案例，重点介绍了 SDN 在数据中心网络和 WAN 网络两个成熟市场的应用。第 7 章介绍了网络虚拟化的发展现状，以及 SDN 与网络虚拟化结合的相关产品。第 8 章作为本书最后一章，梳理了 SDN 对学术界和工业界产生的影响，讨论了我们应该如何应对 SDN 这场变革。书籍附录部分还有《我是一个 SDN 控制器》和《我是一个 SDN 交换机》2 篇拟人化的 SDN 文章供读者阅读。希望这 2 个小故事能让读者对 SDN 有感性的认识。

本书的 1、4、6、8 章由杨泽卫完成，2、3、5、7 章和 2 个技术故事由李呈完成。本书内容是我们学习 SDN 过程中的总结，其目标读者是 SDN 初学者和进阶学习者。我们相信：一本技术书籍的生命力更加体现在后期读者的反馈上，它不仅仅是手边的纸质手册，更像是一个互动交流的学习平台，一个可以让作者和读者一起探讨 SDN 的平台。此外，由于作者水平有限，在书籍内容的编写上难免会有疏漏，观点难免有失偏颇，也恳请读者批评指正。

我们衷心地感谢那些帮助我们完成本书的人们。首先我们要感谢对方，我们一起完成了一本技术书籍。在写作的过程中，我们有过观点探讨时的针锋相对，也有过迷茫懈怠时的相互鼓励，也正因为如此，才能顺利地完成这本书。其次我们要感谢本书的策划编辑张春雨先生，没有他的支持，也就没有本书的诞生。最后我们要感谢我们的家人和师长。感谢在本书写作过程中给予我们启发和帮助的人。谢谢你们！

杨泽卫、李呈

2017年2月5日

读者服务

轻松注册成为博文视点社区用户（www.broadview.com.cn），您即可享受以下服务。

- **提交勘误：**您对书中内容的修改意见可在【提交勘误】处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **与作者交流：**在页面下方【读者评论】处留下您的疑问或观点，与作者和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/31042>

也可扫描下方二维码进入本书页面。



目录

第 1 章 SDN 重塑网络	1
1.1 SDN 是什么	1
1.2 为什么需要 SDN	4
1.3 网络可编程探索之路	6
1.4 SDN 发展历史	10
1.5 SDN 重塑网络	15
1.6 本章小结	16
参考资料	16
第 2 章 SDN 南向协议	21
2.1 SDN 南向协议简介	21
2.2 狭义 SDN 南向协议	23
2.3 广义 SDN 南向协议	30
2.3.1 OF-Config	31
2.3.2 OVSDB	34
2.3.3 NETCONF	36
2.3.4 OpFlex	38
2.3.5 XMPP	39
2.3.6 PCEP	40
2.4 完全可编程南向协议	41
2.4.1 POF	41
2.4.2 P4	46

2.5	SDN 南向协议标准之战.....	53
2.6	本章小结.....	55
	参考资料.....	55
第 3 章	SDN 控制平面	57
3.1	SDN 控制平面简介.....	57
3.2	SDN 开源控制器.....	59
3.2.1	NOX/POX.....	59
3.2.2	Ryu.....	63
3.2.3	Floodlight.....	68
3.2.4	OpenDaylight.....	72
3.2.5	ONOS.....	78
3.3	选择 SDN 控制器.....	85
3.3.1	评价控制器的要素.....	85
3.3.2	选择正确的控制器.....	89
3.4	SDN 控制平面发展趋势.....	91
3.5	本章小结.....	96
	参考资料.....	97
第 4 章	SDN 数据平面	99
4.1	SDN 数据平面简介.....	99
4.2	通用可编程转发模型.....	101
4.2.1	通用硬件模型.....	103
4.2.2	通用处理指令.....	112
4.2.3	小结	115
4.3	探索通用可编程数据平面.....	116
4.4	SDN 数据平面的发展趋势.....	120
4.4.1	发展历史.....	120
4.4.2	白盒交换机.....	123
4.5	本章小结.....	125
	参考资料.....	126
第 5 章	从零开始实践	129
5.1	Mininet 实践.....	129

5.1.1	Mininet 简介	130
5.1.2	Mininet 系统架构	130
5.1.3	Mininet 安装	132
5.1.4	Mininet 示例	134
5.2	Open vSwitch 实践	140
5.2.1	OVS 简介	141
5.2.2	OVS 架构	141
5.2.3	OVS 安装	142
5.2.4	OVS 示例	144
5.3	Ryu 实践	150
5.3.1	Ryu 简介	150
5.3.2	Ryu 架构	151
5.3.3	Ryu 安装	152
5.3.4	Ryu 示例	153
5.4	网络虚拟化平台实践	162
5.4.1	OpenVirteX 简介	162
5.4.2	OpenVirteX 架构	163
5.4.3	OpenVirteX 安装	164
5.4.4	OpenVirteX 示例	165
5.4.5	其他网络虚拟化产品	170
5.5	其他工具	170
5.5.1	Cbench 简介	171
5.5.2	OFTest 简介	173
5.5.3	Wireshark 简介	174
5.5.4	发包工具简介	175
5.6	本章小结	177
	参考资料	177
第 6 章	SDN 应用案例	179
6.1	SDN 在数据中心网络的应用	179
6.1.1	面临的问题	179
6.1.2	现有商业方案	181

6.2	SDN 在 WAN 中的应用.....	187
6.2.1	面临的问题.....	187
6.2.2	现有商业方案.....	189
6.3	其他领域的应用.....	192
6.3.1	SDN 在园区网中的应用.....	193
6.3.2	SDN 在局域网中的应用.....	195
6.3.3	SDN 在蜂窝网络中的应用.....	195
6.4	本章小结.....	197
	参考资料.....	197
第 7 章	SDN 与网络虚拟化.....	200
7.1	网络虚拟化.....	200
7.1.1	为什么需要虚拟化技术.....	201
7.1.2	网络虚拟化.....	203
7.2	SDN 与网络虚拟化.....	208
7.2.1	SDN 与网络虚拟化.....	208
7.2.2	SDN 实现网络虚拟化.....	209
7.3	网络虚拟化产品.....	213
7.3.1	开源产品.....	214
7.3.2	商业产品.....	220
7.4	未来研究方向.....	223
7.5	本章小结.....	225
	参考资料.....	225
第 8 章	SDN 浪潮.....	227
8.1	SDN 对学术界的影响.....	227
8.2	SDN 对工业界的影响.....	231
8.3	如何应对 SDN 的变革.....	234
8.4	SDN 浪潮.....	238
8.5	本章小结.....	238
	参考资料.....	239
附录 A	我是一个 SDN 控制器.....	243
附录 B	我是一个 SDN 交换机.....	250

第 1 章

SDN 重塑网络

自 SDN (Software-Defined Networking, 软件定义网络) 诞生至今已经有七个年头了, 无论在学术界还是产业界, SDN 都引起了广泛的关注。作为一种新的网络体系结构, SDN 将重塑网络行业的竞争格局。所以无论对于学术界的研究人员还是产业界的从业者而言, 我们都需要做好准备, 主动拥抱这场网络的变革。

1.1 SDN 是什么

有人认为, SDN 就是数控分离; 有人认为, SDN 就是 OpenFlow; 还有人认为只要支持软件编程控制的网络就是 SDN。不同的人对 SDN 有着不一样的理解, 正如一千个人眼中有一千个哈姆雷特。为探究 SDN 的准确定义, 我们将以 ONRC^[1] (OpenFlow Network Research Center, 开放网络研究中心) 和 ONF^[2] (Open Networking Foundation, 开放网络基金会) 对 SDN 的定义为切入点, 深入探讨 SDN 的本质, 一层一层揭开 SDN 的神秘面纱, 直到看清 SDN 的庐山真面目。

ONRC 是 SDN 创始人斯坦福大学教授 Nick McKeown^[3] 和加州大学伯克利分校教授 Scott Shenker^[4], 以及大名鼎鼎的 Larry Peterson 教授^[5] 共同创建的研究架构。ONRC 对

SDN 的定义是：“SDN 是一种逻辑集中控制的新网络架构，其关键属性包括：数据平面和控制平面分离；控制平面和数据平面之间有统一的开放接口 OpenFlow。”在 ONRC 的定义中，SDN 的特征表现为数据平面和控制平面分离，拥有逻辑集中式的控制平面，并通过统一而开放的南向接口来实现对网络的控制。ONRC 强调了“数控分离”，逻辑集中式控制和统一、开放的接口。

相比 ONRC 对 SDN 的定义，另一个重要的组织 ONF 对 SDN 定义做出了不同的描述。ONF 是 Nick McKeown 教授和 Scott Shenker 教授联合多家业界厂商发起的非营利性开放组织，其工作的主要内容是推动 SDN 的标准化和商业化进程。ONF 认为：“SDN 是一种支持动态、弹性管理的新型网络体系结构，是实现高带宽、动态网络的理想架构。SDN 将网络的控制平面和数据平面解耦分离，抽象了数据平面网络资源，并支持通过统一的接口对网络直接进行编程控制”。相比之下，ONF 强调了 SDN 对网络资源的抽象能力和可编程能力。

本质上，这两个组织给出的 SDN 定义并没有太大的差别，都强调了 SDN 拥有数据平面和控制平面解耦分离的特点，也都强调了 SDN 支持通过软件编程对网络进行控制的能力。但是 ONRC 更强调数控分离和集中控制等表现形式，而 ONF 则强调抽象和可编程等功能。

从 ONRC 和 ONF 对 SDN 的定义中可以了解到：SDN 不仅重构了网络的系统功能，实现了数控分离，也对网络资源进行了抽象，建立了新的网络抽象模型。SDN 主要有如下三个特征。

（1）网络开放可编程：SDN 建立了新的网络抽象模型，为用户提供了一套完整的通用 API，使用户可以在控制器上编程实现对网络的配置、控制和管理，从而加快网络业务部署的进程。

（2）控制平面与数据平面的分离：此处的分离是指控制平面与数据平面的解耦合。控制平面和数据平面之间不再相互依赖，两者可以独立完成体系结构的演进，类似于计算机工业的 Wintel 模式，双方只需要遵循统一的开放接口进行通信即可。控制平面与数据平面的分离是 SDN 架构区别于传统网络体系结构的重要标志，是网络获得更多可编程能力的架构基础。

（3）逻辑上的集中控制：主要是指对分布式网络状态的集中统一管理。在 SDN 架构中，控制器会担负起收集和管理所有网络状态信息的重任。逻辑集中控制为软件编程定义网络功能提供了架构基础，也为网络自动化管理提供了可能。

因此,只要符合以上三个特征的网络都可以称之为软件定义网络。在这三个特征中,控制平面和数据平面分离为逻辑集中控制创造了条件,逻辑集中控制为开放可编程控制提供了架构基础,而网络开放可编程才是 SDN 的核心特征。

一般来说,SDN 网络体系结构主要包括 SDN 网络应用、北向接口、SDN 控制器、南向接口和 SDN 数据平面共五部分,如图 1-1 所示。

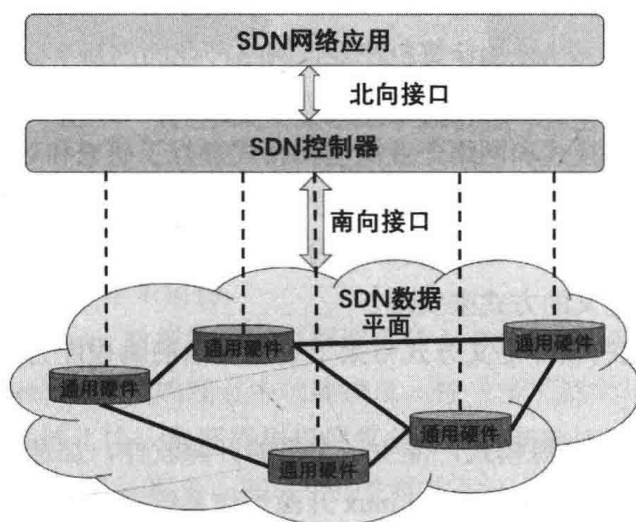


图 1-1 SDN 体系架构图

SDN 网络应用层实现了对应的网络功能应用。这些应用程序通过调用 SDN 控制器的北向接口,实现对网络数据平面设备的配置、管理和控制。

北向接口是 SDN 控制器与网络应用之间的开放接口,它将数据平面资源和状态信息抽象成统一的开放编程接口。

SDN 控制器是 SDN 的大脑,也称作网络操作系统。控制器不仅要通过北向接口给上层网络应用提供不同层次的可编程能力,还要通过南向接口对 SDN 数据平面进行统一配置、管理和控制。

南向接口是 SDN 控制器与数据平面之间的开放接口。SDN 控制器通过南向接口对数据平面进行编程控制,实现数据平面的转发等网络行为。

SDN 数据平面包括基于软件实现的和基于硬件实现的数据平面设备。数据平面设备通过南向接口接收来自控制器的指令,并按照这些指令完成特定的网络数据处理。同时,SDN 数据平面设备也可以通过南向接口给控制器反馈网络配置和运行时的状态信息。

1.2 为什么需要 SDN

了解了 SDN 的定义之后，读者不禁要问：“SDN 为什么会出现？是什么原因使得学术界提出 SDN？而我们又为什么需要 SDN？”这些问题也正是笔者最近两年在推广 SDN 方案时，听到最多的疑问。问题的答案是：我们需要拥有更多可编程能力的网络，来支持快速增长的网络业务需求。

众所周知，相比发展迅速的计算机产业，网络产业的创新十分缓慢。每一个创新都需要等待数年才能完成技术标准化。为了解决这个问题，SDN 创始人 Nick McKeown 教授对计算机产业的创新模式和网络产业的创新模式进行了研究和对比。在分析了计算机产业的创新模式之后，他总结出支撑计算机产业快速创新的如下三个因素。

(1) 计算机工业找到了一个面向计算的通用硬件底层：通用处理器，使得计算机的功能可以通过软件定义的方式来实现。

(2) 计算机功能的软件定义方式带来了更加灵活的编程能力，使得软件应用的种类得到爆炸式的增长。

(3) 计算机软件的开源模式，催生了大量的开源软件，加速了软件开发的进程，推动了整个计算机产业的快速发展，Linux 开源操作系统就是最好的证明。

相比之下，传统的网络设备与上世纪 60 年代的 IBM 大型机类似，网络设备硬件、操作系统和网络应用三部分紧耦合在一起组成一个封闭的系统。这三部分相互依赖，通常隶属于同一家网络设备厂商，每一部分的创新和演进都要求其余部分做出同样的升级。这样的架构严重阻碍了网络创新进程的开展。如果网络产业能像当今计算机产业一样，也具备通用硬件底层、软件定义功能和开源模式三要素，一定能获得更快的创新速度，最终像计算机产业一样取得空前的发展。

正是在这种思路的影响下，McKeown 教授团队提出了一个新的网络体系结构：SDN。在 SDN 架构中，网络的控制平面与数据平面相分离，数据平面将变得更加通用化，变得与计算机通用硬件底层类似，不再需要具体实现各种网络协议的控制逻辑，而只需要接收控制平面的操作指令并执行即可。网络设备的控制逻辑转而由软件实现的 SDN 控制器和 SDN 应用来定义，从而实现网络功能的软件定义化。随着开源 SDN 控制器和开源 SDN 开放接口的出现，网络体系结构也拥有了通用底层硬件、支持软件定义和开源模式三个要素。从传统网络体系结构到 SDN 网络体系结构的演进关系如图 1-2 所示。