



网络操作系统SONiC

原理、技术与实践

人民邮电出版社

SONiC

网络操作系统

原理、技术与实践

张 玮 史慧玲 谭立状 吴迅亮 等◎著



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

SONiC

网络操作系统

原理、技术与实践

张 玮 史慧玲 谭立状 吴迅亮 等◎著

人民邮电出版社
北 京

图书在版编目 (CIP) 数据

网络操作系统 SONiC：原理、技术与实践 / 张玮等著. -- 北京：人民邮电出版社，2024. -- ISBN 978-7-115-65464-9

I. TP316.8

中国国家版本馆 CIP 数据核字第 2024KT1804 号

内 容 提 要

网络操作系统作为网络基础设施的核心，其重要性日益凸显。本书从理论及实践角度，全方位地介绍了 SONiC 网络操作系统的核心技术。本书主要分为 5 个部分。首先，介绍了 SONiC 的起源、技术特点、功能发展、技术优势、厂商实践及标准化，帮助读者全面了解 SONiC。接着，梳理了 SONiC 的系统架构、SAI、Docker 技术、Redis 数据库及其他关键模块，为读者深入理解 SONiC 提供基础知识。然后，探讨了代码仓库结构、编译 SONiC 镜像的流程、安装/部署 SONiC 的方法和常用命令，使读者能够实际操作 SONiC。此外，通过实例介绍了 SONiC 中的二层和三层网络功能，深入剖析了其核心网络协议。最后，通过典型功能测试，帮助读者掌握各种常见路由功能的配置方法。

本书内容丰富，语言通俗易懂，叙述深入浅出，可作为高等院校计算机相关专业网络操作系统课程的教学参考书，也可作为网络工程师、系统管理员、开发者和 IT 运维人员等的技术指导书。

◆ 著 张 玮 史慧玲 谭立状 吴迅亮 等

责任编辑 秦萃青

责任印制 马振武

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号

邮编 100164 电子邮件 315@ptpress.com.cn

网址 <https://www.ptpress.com.cn>

固安县铭成印刷有限公司印刷

◆ 开本：690×970 1/16

印张：15.5

2024 年 12 月第 1 版

字数：278 千字

2024 年 12 月河北第 1 次印刷

定价：149.80 元

读者服务热线：(010)53913866 印装质量热线：(010)81055316

反盗版热线：(010)81055315

广告经营许可证：京东市监广登字 20170147 号



前言

在数字化时代，网络技术需求与发展日新月异。网络操作系统作为网络基础设施的核心，其重要性日益凸显。2016年，微软在开放计算项目全球峰会上首次发布基于 Debian GNU/Linux 的开源网络交换机操作系统——SONiC，通过解耦网络控制面与数据转发面，支持白牌交换机灵活组网，允许用户高效、快速地开发、调试、部署和修复网络软件功能，为构建可扩展、高性能的大规模网络提供无限可能。自 SONiC 发布以来，国内互联网及网络设备厂商积极跟进。例如，2017年，阿里巴巴加入 SONiC，线上部署百万个 SONiC 网络端口，发起并积极推动 SRv6 on SONiC 项目。2021年，浪潮基于 SONiC 研发网络操作系统 Inspur NOS，并不断进行 SONiC 特性增强、可靠性加固、性能优化和场景测试验证。目前，SONiC 作为 Linux 基金会正式项目，已经成为开源网络操作系统的事实标准，被云服务提供商、交换机供应商、专用集成电路（ASIC）芯片制造商及大型企业广泛采纳，成为大规模网络的重要选型技术。

尽管 SONiC 的发展势头迅猛，但在理论和实践方面的资料相对匮乏，很多希望学习、部署和开发 SONiC 的读者面临较大的困难。此外，由于 SONiC 技术新颖，市面上的图书资料缺乏系统性介绍和指导，这限制了 SONiC 在更广泛范围内的应用和推广。基于上述背景，我们决定编写一本全面介绍 SONiC 的技术书，以帮助读者从理论到实践全面掌握构建可扩展、高性能网络的核心技术。本书力求深入浅出，将理论与实践相结合，以满足不同读者的需求。

本书主要关注 SONiC 网络架构的相关内容，共分为 5 章，基本涵盖了 SONiC 的关键内容。第 1 章主要介绍了 SONiC 的起源、技术特点、功能发展、技术优势、厂商实践及标准化，帮助读者对 SONiC 建立全面、初步的认识。第 2 章详细介绍了 SONiC 的系统架构、硬件解耦合的 SAI、软件解耦合的 Docker 技术、数据库驱动的 Redis 数据库及关键模块（如 SwSS 模块和 Syncd 模块），帮助读者理解 SONiC 的关键架构和主要组件。第 3 章介绍了 SONiC 的代码仓库结构、编译 SONiC 镜像的流程、通过 ONIE 安装 SONiC 的方法、通过 GNS3 部署 SONiC



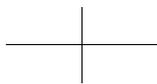
网络操作系统SONiC：原理、技术与实践

的方法，以及 SONiC 的常用命令，帮助读者掌握实际部署和应用 SONiC 的方法。第 4 章通过实例介绍了 SONiC 中的二层网络功能（如 VLAN 和 MAC）、三层网络功能（如静态路由和 OSPF），帮助读者深入理解 SONiC 中的核心网络协议。第 5 章通过 VLAN、VLAN 间路由、RIP、EIGRP、OSPF、BGP 和 RIPng 等典型网络功能测试，帮助读者了解 SONiC 中的核心网络协议和掌握 SONiC 中各种路由功能的配置方法。

本书由张玮、史慧玲组织撰写并统稿，谭立状负责 SONiC 核心组件的研究与撰写工作，谷鹏飞负责 SONiC 搭建与部署的研究与撰写工作，吴迅亮负责全书技术结论的验证和审核工作。第 1 章及第 2 章由史慧玲、张玮执笔，第 3 章的使用案例与部署等工作由谭立状、丁伟、郝昊、王小龙实践并记录，第 4 章由吴迅亮、刘朔晗、张义伟、李明发、李俊豪执笔，第 5 章由谷鹏飞执笔。各位作者在各异的领域中都有丰富的实践经验和深入的理论研究，为本书的撰写提供了有力的支持。为了更好地理解和应用本书中的内容，读者需要具备一定的编程经验和网络基础知识。

本书在撰写过程中，得到了齐鲁工业大学（山东省科学院）计算机科学与技术学部、算力互联网与信息安全教育部重点实验室、国家自然科学基金、山东省自然科学基金的支持及业界专家学者、出版社编辑的广泛帮助，在此一并表示感谢。此外，感谢广大读者的关注，希望本书能为读者提供有价值的参考和帮助。

作者
2024 年 7 月



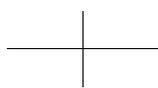
目 录

第 1 章 初识 SONiC	1
1.1 SONiC 概述	1
1.2 技术特点	2
1.3 功能发展	3
1.4 技术优势	4
1.5 厂商实践	6
1.6 SONiC 标准化	7
1.7 本章小结	8
参考文献	9
第 2 章 SONiC 核心组件	10
2.1 SONiC 系统架构	10
2.1.1 系统概述	10
2.1.2 Docker 容器	13
2.1.3 子系统交互	16
2.1.4 消息通信机制	22
2.2 硬件解耦合：SAI	43
2.2.1 SAI 概述	43
2.2.2 SAI 的接口定义	43
2.2.3 数据结构说明及初始化	44
2.2.4 关键组件	47
2.2.5 SAI-ACL 模块	48
2.2.6 SAI 实现	49
2.2.7 Pipeline 定义	50
2.2.8 SAI 使用	50



2.3	软件解耦合： Docker 技术	52
2.3.1	Docker 概述	52
2.3.2	基础命令	52
2.3.3	构建镜像	53
2.3.4	网络模型	54
2.3.5	Docker 原理	54
2.4	SwSS 模块	55
2.4.1	SwSS 概述	55
2.4.2	SwSS 启动	55
2.4.3	*syncd 进程	58
2.4.4	*mgrd 进程	60
2.4.5	Orchagent 概述	61
2.5	Syncd 模块	62
2.5.1	Syncd 概述	62
2.5.2	Syncd 启动	62
2.5.3	Syncd 进程	64
2.6	数据库驱动： Redis 数据库	67
2.6.1	Redis 概述及功能解析	67
2.6.2	以数据库为中心的模型	69
2.6.3	与内核的通信方式	72
2.7	开源路由协议栈（FRRouting）	76
2.8	可编程芯片	79
2.9	服务和 workflow	80
2.9.1	服务分类	81
2.9.2	服务间控制流分类	82
2.10	核心容器	83
2.10.1	数据库容器： Database 容器	84
2.10.2	交换机状态管理容器： SwSS 容器	85
2.10.3	ASIC 管理容器： Syncd 容器	85
2.10.4	各种实现特定功能的容器	86
2.10.5	管理服务容器： mgmt-framework 容器	86
2.10.6	平台监控容器： PMON 容器	87
2.11	本章小结	88
	参考文献	89

第 3 章 SONiC 系统实践	90
3.1 代码仓库	90
3.1.1 核心仓库	90
3.1.2 功能实现仓库	91
3.1.3 工具仓库: sonic-utilities	95
3.1.4 内核补丁: sonic-linux-kernel	95
3.2 编译 SONiC 镜像	96
3.2.1 编译环境搭建	96
3.2.2 编译过程	102
3.3 通过 ONIE 安装 SONiC	111
3.3.1 安装 ONIE	113
3.3.2 安装 SONiC	115
3.3.3 SONiC 镜像升级	118
3.4 通过 GNS3 部署 SONiC	122
3.4.1 安装 GNS3	122
3.4.2 创建网络	128
3.4.3 配置网络	130
3.5 常用命令	133
3.6 本章小结	135
第 4 章 典型网络协议分析	137
4.1 概述	137
4.2 二层网络功能	138
4.2.1 VLAN	138
4.2.2 MAC	149
4.3 三层网络功能	152
4.3.1 静态路由	153
4.3.2 OSPF	159
4.4 网络监控	170
4.4.1 Telemetry 概述	170
4.4.2 Telemetry 相关协议	171
4.4.3 数据源	171
4.4.4 订阅模式	176
4.5 SONiC 无损网络实现	178



4.5.1	RDMA 概述	179
4.5.2	支持 RDMA 的协议	179
4.5.3	无损网络概述	181
4.5.4	DCB 概述	182
4.5.5	ECN 的实现原理	183
4.5.6	PFC 的实现原理	188
4.5.7	PFC Watchdog	191
4.5.8	PFC 死锁	193
4.5.9	DCB 在芯片中的实现	197
4.6	本章小结	206
	参考文献	207
第 5 章	典型功能测试	209
5.1	VLAN	209
5.1.1	VLAN 概述	209
5.1.2	网络拓扑	211
5.1.3	网络配置	211
5.1.4	连通性测试	214
5.2	VLAN 间路由	215
5.2.1	VLAN 间路由概述	215
5.2.2	网络拓扑	215
5.2.3	网络配置	215
5.2.4	连通性测试	217
5.3	RIP	217
5.3.1	RIP 概述	217
5.3.2	网络拓扑	217
5.3.3	网络配置	218
5.3.4	连通性测试	221
5.4	EIGRP	221
5.4.1	EIGRP 概述	221
5.4.2	网络拓扑	221
5.4.3	网络配置	222
5.4.4	连通性测试	224
5.5	OSPF	224

5.5.1	OSPF 概述	224
5.5.2	网络拓扑	225
5.5.3	网络配置	225
5.5.4	连通性测试	228
5.6	BGP	228
5.6.1	BGP 概述	228
5.6.2	网络拓扑	228
5.6.3	网络配置	229
5.6.4	连通性测试	231
5.7	RIPng	231
5.7.1	RIPng 概述	231
5.7.2	网络拓扑	232
5.7.3	网络配置	232
5.7.4	连通性测试	235
5.8	本章小结	235
	参考文献	236