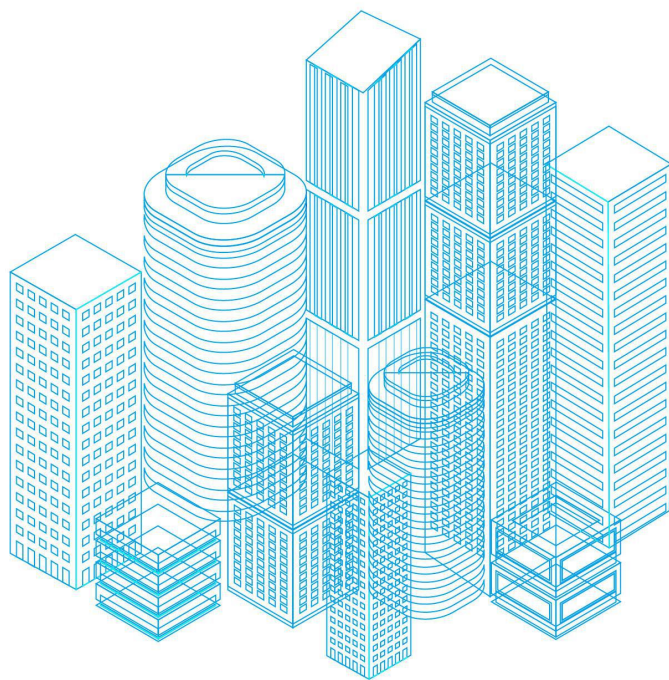




未来网络白皮书

——白盒交换机技术白皮书 (2021版)



第五届未来网络发展大会组委会
2021年6月

版权声明

本白皮书版权属于网络通信与安全紫金山实验室所有并受法律保护，任何个人或是组织在转载、摘编或以其他方式引用本白皮书中的文字、数据、图片或者观点时，应注明“**来源：网络通信与安全紫金山实验室**”。否则将违反中国有关知识产权的相关法律和法规，对此网络通信与安全紫金山实验室有权追究侵权者的相关法律责任。

编写说明

编写单位：

网络通信与安全紫金山实验室、北京邮电大学、江苏省未来网络创新研究院

参与单位：（排序不分先后）

中国联通研究院、中国联合网络通信有限公司江苏省分公司、新华三集团、东南大学、钰登科技股份有限公司、南京华智达网络技术有限公司、星融元数据技术有限公司、瞻博网络

主要编写人员：（排序不分先后）

汪硕、沈洋、吴斌伟、朱华兴、陈平平、李力、张成林、吴芄、周运文、谭炜骞、陈娟娟、彭开来、王佳森、王严严、李振红、吴生三、欧俊明、杨军、宋梅、朱敏晓、林萱、胡茂庐、陈维、王雪、万欢、尹川、徐博华、庞冉、张学茹、井有浩、唐卫君、毕诗章、樊成

前 言

随着互联网技术、移动通信技术的不断演进和发展，网络与社会生活、工业生产逐渐融合，未来新型业务将不断涌现，例如远程医疗、自动驾驶、智能制造、智慧城市、全息通信、虚拟现实等。当前，封闭、不可编程的传统网络架构难以支撑未来网络业务对网络协议与功能的可编程、确定性、定制化、高性能需求。同时，伴随 SDN（Software Defined Network，软件定义网络）、NFV（Network Functions Virtualization，网络功能虚拟化）技术的出现，云计算、云原生等云技术飞速发展，网络开始与云融合，提供更快迭代速率、更开放的管控能力，以及更灵活的业务部署能力。白盒交换机突破传统交换机软硬件的一体化设计，采用开放的设备架构，解耦网络底层硬件与上层网络功能或协议，支持需求的快速迭代，提升了设备的可编程性、灵活性。上述特征契合了新型业务和网络发展的痛点需求，有望为网络带来全新的改变，对我国未来网络建设和研究具有重要意义。

本白皮书梳理了白盒交换机的概念和产生背景，系统调研了白盒交换机的发展现状和未来趋势，对白盒交换机的关键技术、设备架构、应用场景和典型案例进行了系统性的跟踪和调研，为我国产业界白盒化提供了参考。

本白皮书共分为七个章节。第一章首先介绍了白盒交换机出现的背景；第二章调研了白盒交换机的发展历史和未来趋势，并从开源生态和产业生态两个方面阐述了白盒交换机的发展现状；第三章阐述了

白盒交换机的若干关键技术，包括软硬件解耦、网络可编程技术、硬件加速技术和白盒安全技术等；第四章系统性的介绍了目前主流白盒交换机的技术架构，包括底层的硬件转发层、芯片接口层、基础软件平台层、网络操作系统和协议层，通过开放的架构、软硬件解耦和分层协作的思想，白盒交换机的可编程性和灵活性得到了巨大的提升；第五章枚举了若干可能的应用场景，包括数据中心、新型承载网和5G云；为推动白盒交换机技术落地，加快产业发展和生态建设，第六章介绍了若干白盒交换机实际应用案例；最后第七章提出我国白盒交换机发展中将应对的挑战并给出应对建议。

第五届未来网络发展大会

目 录

前 言.....	I
目 录.....	III
图 目 录.....	VI
表 目 录.....	VIII
一、 背景.....	1
二、 发展现状及趋势.....	3
2.1 发展历程.....	3
2.2 开源生态.....	5
2.2.1 开放计算项目.....	5
2.2.2 开放网络基金会.....	6
2.2.3 电信基础设施项目.....	7
2.2.4 开源数据中心委员会.....	8
2.3 产业生态.....	9
2.3.1 设备提供商.....	9
2.3.2 云服务商.....	12
2.3.3 电信运营商.....	16
2.4 未来趋势.....	19
三、 关键技术.....	23
3.1 软硬件解耦技术.....	23
3.2 网络可编程技术.....	25
3.2.1 控制面可编程能力.....	26
3.2.2 数据面可编程能力.....	27

3.3 硬件加速技术.....	30
3.4 白盒安全技术.....	32
四、 设备架构.....	34
4.1 硬件转发层.....	34
4.1.1 BAREFOOT TOFINO 系列芯片.....	35
4.1.2 BROADCOM 可编程系列芯片.....	36
4.1.3 CISCO SILICON ONE 系列芯片.....	38
4.2 芯片接口层.....	39
4.2.1 OPENNSL/OF-DPA.....	39
4.2.2 SAI.....	41
4.3 基础软件平台层.....	43
4.3.1 ONIE.....	43
4.3.2 ONL.....	45
4.4 网络操作系统和协议层.....	47
4.4.1 SONIC.....	47
4.4.2 UNINOS.....	48
4.4.3 DANOS.....	50
4.4.4 STRATUM.....	51
4.4.5 DENT.....	52
4.4.6 总结与对比.....	53
五、 应用场景.....	54
5.1 数据中心场景.....	54
5.2 新型承载网场景.....	56
5.3 5G 云场景.....	59
六、 典型案例.....	61
6.1 数据中心案例.....	61

6.1.1 解决方案.....	61
6.1.2 部署效果.....	62
6.2 校园网机房案例.....	62
6.2.1 解决方案.....	63
6.2.2 部署效果.....	64
6.3 政务网机房案例.....	64
6.3.1 解决方案.....	64
6.3.2 部署效果.....	66
七、 挑战与建议.....	67
7.1 挑战.....	67
7.2 建议.....	69
附录 A：术语与缩略语.....	71
参考文献.....	73

图 目 录

图 2-1 白盒交换机发展历程.....	3
图 2-2 网络闭环控制.....	21
图 3-1 AT&T 白盒交换机生态系统.....	24
图 3-2 控制器体系架构.....	27
图 3-3 传统交换芯片与可编程交换芯片.....	28
图 3-4 传统交换机架构.....	31
图 3-5 服务器+智能网卡型架构.....	31
图 3-6 服务器+交换机+加速卡型架构.....	32
图 4-1 白盒交换机架构.....	34
图 4-2 PISA 交换机架构.....	35
图 4-3 Trident4 芯片架构图.....	37
图 4-4 Silicon One Q200/G100 芯片结构图.....	38
图 4-5 OpenNSL 架构.....	40
图 4-6 OpenFlow 到 OF-DPA 的转换.....	40
图 4-7 交换机抽象接口.....	42
图 4-8 SAI 对不同的 ASIC SDK 进行抽象.....	43
图 4-9 ONL 构成.....	46
图 4-10 SONiC 系统架构.....	48
图 4-11 UniNOS 架构.....	49

图 4-12 DNOS 功能层和组件.....	50
图 4-13 Stratum 系统架构.....	51
图 4-14 DENT 架构.....	52
图 5-1 Spine-Leaf 架构.....	55
图 5-2 白盒化 IP 承载网网络架构.....	56
图 5-3 5G 移动回传和专线业务场景.....	57
图 5-4 固网宽带业务场景.....	58
图 5-5 5G 通信云白盒化场景.....	59
图 5-6 5G 边缘云白盒化场景.....	60
图 6-1 互联网企业数据中心部署示意图.....	61
图 6-2 校园网机房部署示意图.....	63
图 6-3 政务网机房案例部署示意图.....	65

第五届未来网络大会白皮书

表 目 录

表 4-1 Tofino 芯片比较.....	36
表 4-2 Broadcom 芯片比较.....	36
表 4-3 EEPROM 模块示例.....	44
表 4-4 NOS 总结与对比.....	53

第五届未来网络发展大会白皮书

一、背景

近十年来，互联网技术、移动通信技术、计算机技术快速发展，网络开始逐渐融入生活的方方面面，用户数量、网络流量、网络规模每年呈几何级数快速增长，网络所承载的业务也开始变得丰富多样。根据预测，至 2030 年，未来网络将需要承载远程医疗、自动驾驶、智能制造、智慧城市、全息通信、虚拟现实等新型复杂业务。这些业务要求网络能够支持高带宽、低时延、低抖动等服务质量指标，支撑网络协议与网络功能的可编程、确定性、定制化、高性能等需求[1]。另外，面向云计算、云原生等云技术飞速发展的现状，网络需要与云融合，提供更快迭代速率（周级别）、更开放的管控能力，以及更灵活的业务部署能力。然而，当前网络依然采用封闭的、设备商垄断的专用网络设备，只能为网络服务提供商提供指定的网络功能，严重阻碍着网络技术的发展，越来越不能适应未来灵活多变网络业务的态势。

白盒交换机是一种软硬件解耦的开放网络设备，与传统软硬一体的封闭交换机相比具有诸多优势[2]。首先，白盒交换机采用开放的设备架构和软硬解耦思想，可以根据业务需求，按需定制底层硬件和上层软件，相比传统交换机软硬件捆绑购买、垄断使用，能够显著降低交换机的购置成本。另外，在软件功能方面，可以基于开源软件进行二次开发，降低开发周期和成本。其次，白盒交换机支持硬件数据

面可编程和软件容器化部署，通过软件定义的方式定制数据面的转发逻辑，还充分利用现代云计算技术，对网络功能进行快速升级迭代，提升网络的灵活性、敏捷性、确定性，优化网络性能，满足复杂的业务需求。另外，借助容器化部署，能统一简化管理运维，降低网络的运维成本。最后，交换机白盒化已得到了芯片厂商、设备提供商、云服务商、电信运营商等交换机上下游企业的一致认同，可以联动白盒开源生态和产业生态的发展，形成繁荣的白盒化网络生态，最终能够促进网络的不断革新演进，解决当前业务问题、满足未来网络需求。

第五届未来网络发展大会

二、 发展现状及趋势

目前，从研制商用可编程芯片到白盒硬件设备的标准化，从统一芯片接口到开发开源的交换操作系统，白盒交换机已经形成具备产业化能力的网络生态。本章首先简述白盒交换机的发展历程，然后从开源生态和产业生态两个方面对白盒交换机的现状进行介绍，最后阐述白盒交换机相关的未来发展趋势。

2.1 发展历程

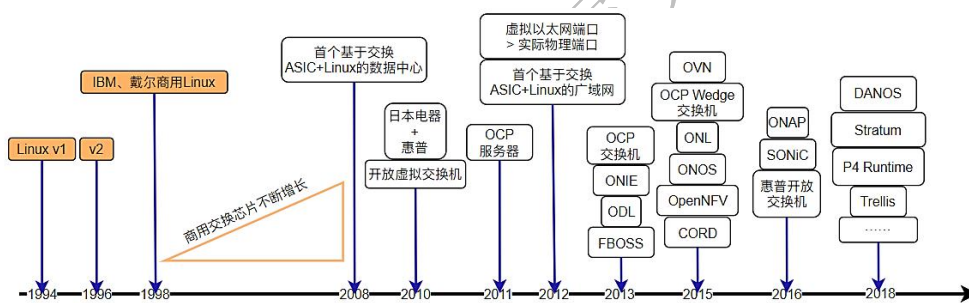


图 2-1 白盒交换机发展历程

如图 2-1 所示，白盒交换机在过去三十年间得到了快速发展，其中开放网络基金会（Open Networking Foundation, ONF）、Linux 基金会、开放计算项目（Open Compute Project, OCP）、电信基础设施项目（Telecom Infra Project, TIP）等开源组织做出了重要贡献。首先，Linux 1.0 版本于 1994 年正式发布，2 年后 2.0 版本正式更新，提供了网络协议/功能控制的开源框架。用户可根据自己的需求，通过修改内核代码，对网络功能与协议进行修改和定制。受此鼓舞，1998 年 IBM、Compaq、Dell 等公司陆续开始商用 Linux 系统，其网络技

术及相关生态开始快速发展。2008年，Linux开始尝试与交换芯片结合，在数据中心场景中提供大容量、高带宽的域内数据传输服务。为进一步推动Linux交换机商业发展，2010年日本电器（NEC）和惠普（HP）着手研究交换机软件化技术，推出基于OVS（OpenVSwitch）的开放软件交换机，网络的资源和能力得到前所未有的释放，网络运营开始走向自动化、智能化[3]。2011年，基于交换机软件化技术，OCP等组织开始关注交换机虚拟化技术，并于2013年开启了对交换机硬件白盒化的标准化工作，推出了ONIE（Open Network Install Environment）开放安装环境、FBOSS（Facebook Open Switching System）设备管理软件以及ODL（OpenDaylight）控制器标准文档，在SDN和白盒交换机领域取得了重大突破。2015年，OCP成功推出第一款白盒交换机Wedge[2]。与此同时，OVN（Open Virtual Network）虚拟化SDN网络、ONL（Open Network Linux）操作系统、ONOS（Open Network Operating System）控制器以及电信领域OpenNFV、CORD（Central Office Rearchitected as a Datacenter）等虚拟化、白盒化项目也相继兴起。2016年至今，白盒设备、软件操作系统、网络自动化等技术已得到蓬勃的发展。微软推出的SONiC（Software for Open Networking in the Cloud）、惠普推出的OpenSwitch、AT&T推出的DANOS（Disaggregated Network Operating System）以及谷歌面向NG-SDN（Next Generation SDN）推出的Stratum，开源交换机操作系统层出不穷。同时，ONAP（Open Network Automation Platform）、

P4Runtime 接口、Trellis 等网络管控解决方案也呼之欲出，白盒交换机相关的网络技术空前繁荣。

2.2 开源生态

白盒交换机开源生态主要围绕几个国内外的开源组织：一是开放计算项目，负责白盒交换机硬件标准的制定；二是开放网络基金会，推进白盒交换机中 SDN 相关技术的发展与落地；三是电信基础设施项目，探索利用白盒交换机技术来改变传统的构建和部署电信网络基础设施的新方法；四是开源数据中心委员会，联合国内机构围绕数据中心基础设备进行开放、合作、创新、共赢的发展。

2.2.1 开放计算项目

开放计算项目（OCP）是由 Facebook 联合英特尔、Rackspace、高盛及 Andy Bechtolsheim 在 2011 年为共享开源设计而推出的一个开放硬件项目，目前已成为一个快速发展的全球合作社区。OCP 专注于重新设计硬件技术，使其更加高效、灵活和可扩展，以支持不断增长的计算基础设施需求。OCP 为个人和组织提供与他人分享知识产权的体系架构，通过开源硬件和软件相结合，促进服务、存储和数据中心技术的开放与普及。OCP 基金会在白盒交换机发展方面的主要工作组有：开放计算网络、高性能计算（High Performance Computing, HPC）、安全、电信等，每个工作组孵化一个或多个项目。

在 2019 年 3 月的 OCP 全球峰会上，Submer 公司推出了全新的 SmartPodX 平台，以及适用于高性能、超级计算和超大规模基础设施的 OCP 规范；同年，华为宣布与 OCP 展开合作，将在其最新的全球数据中心采用 OCP 的 OpenRack，以降低总体购置成本并提高规模计算领域的能效。2019 年 6 月，首届 OCP China Day 在北京正式开启，来自 Facebook、LinkedIn、Intel、微软、百度、腾讯、阿里、诺基亚、中国移动、浪潮等资深技术专家，分享了人工智能、边缘计算、OpenRack、OpenRMC、SONiC、OAM（Operation Administration and Maintenance）等前沿技术的相关进展。2020 年 5 月 12 日，以“OpenForAll”为主题的 2020 年 OCP 全球峰会在线举行，会议展示了多款基于开放计算的下一代数据中心解决方案，加速开放计算走进传统行业。

2.2.2 开放网络基金会

开放网络基金会（ONF）是 SDN 主要提出者 Nick McKeown 及 Scott Shenker 于 2011 年创立的网络领域开源组织，旨在**推进 SDN 的发展与落地**，是 SDN 领域公认的领导者和标准承载者。自成立以来，ONF 已成功推动 SDN 从前景不明到成为被运营商、设备商、服务提供商普遍接受的下一代网络技术。同时 ONF 也在积极探索基于新的转发面编程语言 P4（Programming Protocol-independent Packet Processors）的数据面开发技术，促进软件定义网络以及交换机数据

面可编程的发展。

2019年9月，ONF在下一代SDN计划中取得了重大进展，开源Google研发的Stratum交换机操作系统。Stratum是一个独立于芯片的平台，旨在使网络运营商轻松集成众多供应商的新设备，从而能够实时扩展和升级其基础架构。2019年10月，中国联通联合ONF举办了“中国联通&ONF开源项目启动会”，来自国内外30多家厂商的70余位专家出席会议。会议分别围绕SD-RAN、MAEC、白盒交换机三个项目，展开开放式讨论，明确了下一阶段的研究目标以及合作意向。截至2021年5月，ONF已拥有5个平台：NG-SDN、ONOS、Mininet、CORD、Aether，支持7个开源项目：STRATUM、TRELLIS、P4、XOS、ODTN、OTCC、INFORMATION MODELING & TOOLING，分别致力于SDN/NFV、下一代移动网络、宽带接入、下一代WAN（Wide Area Network）等方面。

2.2.3 电信基础设施项目

电信基础设施项目（Telecom Infra, TIP）是Facebook于2016年主导成立的一个电信领域开放组织，旨在通过共同合作发展新技术，改变传统的构建和部署电信网络基础设施的方法。TIP有500多个成员，包括运营商、基础设施提供商、系统集成商以及其他各类科技公司。Facebook、英特尔和诺基亚等TIP成员承诺提供一套初始参考设计，而其他成员如运营商Deutsche Telekom和SK Telecom将帮助定

义和部署符合他们需要的技术。

在电信领域，TIP 以无线接入为主要应用场景，对整个运营商的网络进行了端到端的覆盖，TIP 的成立使得白盒在运营商领域得到高度的关注。AT&T 2016 年首先宣布 CPE 的白盒计划，后续又宣布计划在未来大量采用白盒来构建 Backhaul 网络，多个欧洲 Tier 1 运营商提出支持 OCP 标准的硬件，如今各国运营商都在积极推动白盒化网络，寻求与设备商解耦的途径。

2.2.4 开源数据中心委员会

开源数据中心委员会（Open Data Center Comittee, ODCC）在中国通信标准化协会指导下，以开放、合作、创新、共赢为宗旨，围绕服务器、数据中心设施、网络、新技术与测试、边缘计算、智能监控与管理等内容进行发展。当前，ODCC 的成员包括腾讯、阿里、百度、中国电信、中国移动、中国信通院、京东和美团等超 100 家会员单位，覆盖数据中心行业的上下游企业。ODCC 定期举办全会、工作组会议、沙龙和研讨会，每年都会发布很多技术白皮书和各种测试标准。

2017 年 8 月，ODCC 网络工作组成员阿里、腾讯、百度、移动、联通、京东、中国电信以及中国信息通信研究院等联合发起了凤凰项目，以 SONiC 开源社区为依托，旨在打造一个开源网络操作系统发行版。2018 年 9 月，凤凰项目发布 1.0 版本，并通过了基本网络功能、访问控制和服务质量、压力、软硬件兼容等测试评估。

2.3 产业生态

在白盒交换机产业生态里，从上游的设备提供商到下游的云服务商、电信运营商，已经在一定规模上形成完整的产业生态链。设备供应商主要包括思科以及新华三集团等，他们面向市场提供更具开放性的类白盒设备解决方案；云服务商主要包括谷歌、微软、阿里巴巴、腾讯等，他们纷纷开始研究白盒交换机操作系统，并借之推新去陈；电信运营商主要包括美国 AT&T、中国移动、中国联通、中国电信等，他们利用白盒交换机进行业务转型和网络重构。

2.3.1 设备提供商

(1) 思科

思科是全球最大的网络设备提供商，近年来其一直围绕 SDN 进行产业发展。思科是业界知名的 SDN 产品提供商，该公司推出的以应用为中心的基础设施（ACI）是业界最初的 SDN 产品之一，主要面向 SDN 的数据中心应用场景。2017 年发布了 ACI 3.0 版本，能够统一管理不同地区的多云环境，为数据中心互连开辟了新的途径。思科也在 SDN 的重要应用场景——WAN 领域研制出名为 IWAN 的产品，在 2017 年收购初创公司 Viptela 后，思科对 IWAN 进行整合，提供新的 SD-WAN（Software-Defined WAN）服务。2019 年 2 月，Verizon 和思科宣布建立合作伙伴关系，利用 Verizon 未来的 5G 网络构建移动 SD-WAN 服务。这是思科第一次在 SD-WAN 上引入移动设备组件。

面向下一个十年的未来网络技术，思科开始布局白盒交换机发展。思科未来互联网战略抛弃了过去种种思维定势，将历史负担清零，对业务进行颠覆性创新。具体包括三大关键技术领域：芯片、光模块和软件。这三个领域的创新不是孤立的，而是会通过各种方式进行融合，形成一个由多个网络层面构成的、一致的多域系统，最终打造成未来互联网的基础设施。在 2019 年 12 月的“Internet for the Future”未来互联网全球发布会上，思科公布了此战略的四个方向：Cisco Silicon One™、面向 400G 和更高速率的光器件、思科 8000 系列平台、灵活的商业模式。目前思科已发布 Cisco Silicon One™ Q100 芯片、思科 8200 系列路由器、思科 8800 系列路由器、光模块 QSFP-DD、IOS XR 7 操作系统等代表性产品。这项战略将重塑互联网经济，全面满足未来数字化应用的需求，并使客户能够使用更加简单、灵活、经济、高效的网络来运营业务。

(2) 新华三集团

新华三集团是国内知名的白盒交换机及网络解决方案提供商，近年来致力于利用白盒交换机助力运营商业务转型和网络升级。2015 年，新华三集团发布了 SD-WAN 解决方案 AD-SDN1.0。2016 年 9 月，新华三集团发布了“新网络、新 IT、新价值”运营商市场战略，致力于解决运营商行业的网络重构挑战。2017 年 11 月，新华三集团在南京发布了 AD-WAN 2.0，通过 SDN 的架构和思想，利用白盒交换机搭

建了一个开放的网络平台，真正实现以业务驱动网络，让静态的网络适应动态的业务需求，真正让网络“动”起来，满足用户实际的业务需求。2020年1月14日至2020年1月19日，SD-WAN项目组通过对新华三集团的AD-WAN解决方案进行全面、严格的测试，全面验证了该方案快速、灵活、高效、智能的软件定义广域网能力。

(3) 星融元数据技术有限公司

星融元数据技术有限公司是新一代开放云网架构解决方案提供商，致力于用开源OS+白盒硬件模式为云计算构建开放的基础网络。2017年，基于微软推出的SONiC，星融开发了AsterNOS开源网络操作系统，该系统能够运行在多种网络硬件芯片厂商的硬件平台之上，基于星融的云网络，用户能够完全按照自己的规划选择合适的硬件平台。2017年至今，星融陆续为开源网络操作系统开发了最合适的硬件生态，主要涵盖组成云数据中心网络的硬件产品--云数据中心交换机、组成云数据中心可视化分析的硬件产品--云网流量采集与分发交换机、以及面向开放网络与计算融合的新一代智能化硬件平台，该系列平台有机融合可编程计算与可编程网络，是打造“算力网络”或“算网融合”支撑基础。在过去的2年里，搭载星融AsterNOS的交换机商业发货量超过6000台，其中有超过5万个100G端口，其客户包括国内外运营商、公有云、网络安全和行业私有云行业等。

(4) 瞻博网络

瞻博网络（Juniper Networks）长期致力于交换机的软硬件解耦，是第一家将数据平面和控制平面分离的路由供应商。基于开放化的、模块化的体系架构，瞻博（Juniper）硬件可支持包括 SONiC 在内的开源网络操作系统，允许客户能够在简化操作的同时，更快地进行创新。同时，针对客户特定应用场景，瞻博网络还做了特定优化。例如，在多芯片机箱式 PTX10000 系列路由器上定制化 SONiC，从而使得使网络运营商可以在 MAN/WAN 核心和数据中心互连场景中使用 SONiC。除了开放化的设备系统，瞻博网络着重强调第三方组件的兼容性，全面支持 SONiC 端到端生态组件，包含芯片、交换平台、系统增强以及路由容器等。其中，瞻博网络的容器化路由协议守护程序（cRPD）可支撑高级 BGP 特性和基于策略的路由（包括 MPLS，SR/SR-TE，L2VPN/L3VPN，EVPN 等能力）。在轻量级部署场景中，还可提供遥测、自动化、可编程性和更快的表项下发性能。

2.3.2 云服务商

(1) 谷歌

谷歌在利用自研白盒交换机探索软件定义网络技术上一一直处于最前沿。其在 SDN 的研发与实践上大致可以分为四个方向：（1）部署全球 SDN 广域网络 B4。2012 年 Google 开始基于自研设备 Saturn（土星），部署全球 SDN 广域网络 B4。2013 年 8 月，发表 B4 SDN

控制器白皮书《B4: Experience with a Globally-Deployed Software Defined WAN》。2018 年发表《B4 and After: Managing Hierarchy, Partitioning, and Asymmetry for Availability and Scale in Google Software-Defined WAN》，描述了 B4 的演进，并更新了自研设备 Stargate（星关）和层次化的 TE 控制器。B4 的实地部署由 2012 年的全球 12 个站点，增加到 2018 年 1 月的 33 个站点，规模大大增加。

(2) 构建网络功能虚拟化 (NFV) 堆栈 Andromeda。2014 年谷歌推出了 Andromeda，构成了 Google Cloud 的基础。其主要目标是使 Google 网络能够为多租赁云提供更大的容量。(3) 构建数据中心互联架构 Jupiter。Jupiter 用于处理单个数据中心内的流量，是一个能够支持超过 10 万台服务器规模的数据中心互联架构。2015 年 Google 通过 SDN 的途径来构建 Jupiter。(4) 构建公共 SDN 互联网 Espresso。2017 年的 ONS 大会上谷歌发布了 Espresso。Espresso 是用于公共互联网的 SDN，主要负责与互联网服务提供商的对等连接。Espresso 将 SDN 扩展到谷歌网络的边缘，连接到全球其他网络，同时根据网络连接的实时测量数据，动态智能化地为个人用户提供服务。

除了自身研发部署白盒交换机外，谷歌还积极参与开源社区，为很多开源项目做出贡献。例如 2018 年 3 月，开放网络基金会 (ONF) 发布了下一代软件定义网络 (NG-SDN) 接口战略，并在谷歌的支持下推出了 Stratum 项目，该项目的大部分工作经验是基于谷歌早期部署 SDN 时发现的 OpenFlow 的局限性。Stratum 旨在为白盒交换机构

建一个开放的、最小化的、生产就绪的发行版，提供完整的白盒交换机解决方案，真正实现 SDN 的“软件定义”承诺。目前，Stratum 公开了一组下一代 SDN 接口，包括 P4、P4Runtime、gNMI/OpenConfig 和 gNOI，实现了转发设备的可互换性和转发行为的可编程性。

(2) 微软

近年来，基于云数据中心部署自研白盒交换机的经验，微软在开源社区上做出了诸多贡献。2015 年 3 月，在 OCP 峰会上，微软向 OCP 贡献了 SAI（Switch Abstraction Interface，交换机抽象接口），并于同年 7 月获得接纳。SAI 对于白盒交换机和 SDN 转控分离的意义十分重大，目前已经得到了博通、Mellanox、Cavium 等主要 ASIC（Application Specific Integrated Circuit）厂商的支持。在 2016 年度开放计算项目（OCP）峰会上，微软正式推出基于 Debian GNU/Linux 的网络交换机操作系统 SONiC。SONiC 是微软发布的开源交换机操作系统，是一个成熟的 SDN 软件平台，实现了数据控制面与转发面的分离。2017 年微软对 SONiC 的性能进行了大幅升级，全面支持 IDV，并且融合了更多的容器特性。2018 年微软通过加入 ConfigDB 功能，支持管理者进行控制和配置。同时，SONiC 还会在深度学习等 AI 特性上做更多的尝试。目前 SONiC 已经得到了业界的广泛认可和采用，众多厂商如 Barefoot、博通、领英、Cavium 等都增加了对 SONiC 的支持，中国的开放数据中心委员会（ODCC）旗下的凤凰项目也是基

于 SONiC 进行开发。

(3) 阿里

阿里是国内顶级的云数据中心服务商，近年来一直积极参与国内白盒设备操作系统的研发。阿里参与开放数据中心委员会并主导了多个核心项目。其中 2017 年 8 月发起的凤凰项目是 ODCC 网络工作组发起的开源网络 OS 发行版项目。其目标是推动“白盒+开源 OS”的网络生态，促进中国开放网络和 SDN 网络的进一步发展。凤凰项目依托 SONiC 开源社区，选择社区中稳定的 SONiC 版本软件，将 Base OS、SAI、平台驱动等共同打包制作成可直接加载运行在白盒交换机的网络 OS 映像文件，即凤凰发行版。另外，凤凰项目也会在软硬件兼容性测试认证、运维管理体系等方面进行探索。

(4) 腾讯

腾讯在开放网络和 SDN 起步较早，近年来一直探索利用白盒交换机进行网络的升级重构。2015 年，腾讯开始在广域 DCI 网络试点 SDN 路径集中计算架构。经过多年积累与演进，在 2017 年 Q4，腾讯成功在 DCI 网络上线了全球首个基于纯交换芯片平台的 SDN based SR-TE 网络架构。该架构基于 SR（Segment Routing）技术和 SDN 思想，率先实现了在数十万服务器级别的多个 IDC 园区间进行通讯，实现全局视角的、多维度的、可定义约束条件的智能调度。此外，在公网流量调度方面，利用 SDN 的思想构建了基于业务指标、网络质

量、成本等多维度的公网流量调度系统，实现了应用驱动的公网流量调度平台。

同时，腾讯也开始利用白盒交换机的开放性在 SD-WAN 领域有所作为。2019 年，腾讯云在 Techo 开发者大会上国内首发云原生 5G SD-WAN 产品。腾讯云 SD-WAN 即插即用，自动化部署，极大地降低了安装使用成本，使网络扩展更加高效迅速。腾讯云首批接入 SD-WAN 的硬件产品有两款：Edge-100 和 Edge-1000。其中，Edge-100 应用于海量分支、智慧门店等组网场景，Edge-1000 应用于 IDC 数据中心、企业总部等核心节点大带宽内网互联。同时，Edge-1000 具备 5G 能力，将是国内首个云原生 5G 的 SD-WAN 网关。腾讯云 SD-WAN 核心优势在于 SDN SR-TE 技术的 TB 级全球互联网络，为企业提供高质量、大带宽、低时延的网络互联，实现云网融合。

2.3.3 电信运营商

(1) AT&T

近年来，AT&T 一直致力于利用白盒交换机实现软件化的云架构开放网络。继 2009 年发布了 Domain1.0 计划后，于 2013 年，AT&T 发布了 Domain2.0 计划。计划通过 SDN/NFV 技术将网络基础设施从“以硬件为中心”转向“以软件为中心”，实现基于云架构的开放网络。在网络软件化领域 AT&T 主要面向如下两个领域展开布局：(1) 针对网络云服务的 ONAP 操作系统项目。2017 年，AT&T 将其内部运行

几年的 Open ECOMP 项目与中国移动等主导的 Open-O 项目合并成为 ONAP 项目,并于 2017 年下半年发布了 ONAP 首个版本 Amsterdam,于 2018 年 6 月发布 Beijing 版本。同年 12 月,发布 Casablanca 版本。目前 ONAP 所涉及的运营商拥有的移动用户占全球移动用户的 70% 以上;(2) 针对虚拟网络环境的白盒服务操作系统项目 DANOS 和白盒设备。2017 年 6 月,AT&T 收购了 Vyatta 网络操作系统及相关资产。随后于 11 月,推出白盒操作系统(即分布式网络操作系统,dNOS)。在 2018 年开放网络峰会(ONS)上,AT&T 宣布将其开源 dNOS 托管给 Linux 基金会,并且将项目名称改为 DANOS。2018 年 12 月,AT&T 开始针对蜂窝基站白盒路由器开展大规模生产测试,并于 2019 年开放网络峰会(ONS),正式发布了白盒化基站网关路由器产品。

(2) 中国移动

中国移动自 2015 年在世界移动大会上提出 Novonet 以来,一直在重点关注新的架构、新的运营与新的服务,尝试利用白盒交换机所带来的网络开放性进行业务转型和网络重构。2016 年中国移动加入软件定义网络平台开源组织 OpenDaylight,希望实现网络环境的标准化、定制化和网络自动化。2016 年,中国移动开始建设最大规模的 SDN 云数据中心——呼哈私有云项目,并在之后陆续建设了多个不同地区的 SDN 云数据中心。2018 年 2 月 27 日,在西班牙巴塞罗那举行的 2018 GTI 国际产业峰会上,中国移动、美国 AT&T、德国电

信、日本 NTT DOCOMO 以及法国 Orange 等五家电信运营企业宣布联合成立 O-RAN 联盟，旨在将下一代无线通信网络的开放性提升到新的水平。目前 O-RAN 联盟已有 23 个运营商成员和 143 个供应商成员。2019 年 11 月，中国移动还携手多家运营商以及供应商，一起成立了开放测试和集成中心（OTIC），旨在通过对开放解耦的无线接入网功能组件进行验证、集成和测试，促进原始设备制造商的产品和其他开源产品及解决方案在功能上符合 O-RAN 联盟的规范，最终实现多功能组件间即插即用、灵活组合搭配的理想架构。

(3) 中国联通

近年来，中国联通一直在尝试利用白盒化网络设备进行网络转型升级。2015 年 9 月，中国联通对外发布了“面向云服务的泛在宽带弹性网络”（CUBE-Net 2.0），倡导云管协同的网络发展理念，致力于推动基于 CUBE-Net 2.0 的网络转型和业务创新。2016 年，中国联通启动了产业互联网发展战略，旨在从泛在连接、弹性网络、云网协同、大数据等产业需求入手，打造基于 SDN 的产业互联网基础设施 CUII。2017 年，中国联通将基于分布全国的综合承载传送网（UTN）网络完成云接入网的 SDN 化改造，并在转发面和控制面实现与 A 网的端到端互通，支持企业分支的敏捷专线需求。2018 年，在云承载网络（云互联和云接入）敏捷的基础上，进一步通过协同提供云业务与网络专线的一体化服务。2018 年 3 月，发布了“云联网”等多个面向政企

客户和 OTT 的云网协同新产品，率先开启了国内运营商提供云网一体化服务的进程，推进网络重构，加快 5G 商用，转型成为数字化服务提供商。

(4) 中国电信

面向 CTNet 2025 新型网络架构，中国电信一直致力于充分利用白盒优势。中国电信在 2016 年 7 月发布了《CTNet 2025 网络架构白皮书》，全面启动网络智能化重构。阶段性目标是到 2020 年，云网协同能力达到一定规模，基础设施云化基本部署；到 2025 年，全网实现智能化、集约化和开放性。2018 年 3 月，中国电信已完成了新一代网络系统的顶层设计。截至目前，中国电信 CTNet 2025 网络重构的实践进展主要分为三大层面：（1）SDN 方面，中国电信面向中小企业发布了随选智能专线产品和解决方案。同时，推出了面向家庭客户的 SDN 智能网关，在多省进行商用试验部署。（2）NFV 方面，中国电信发布了三层解耦的 MANO 和 NFVI 企业标准，完成 vBRAS、vIMS 实验室和现网三层解耦测试。同时，与自研 NFVO 结合，在多省进行试验部署，以承载 VoLTE、VoIP、ITMS 等类型业务。（3）运营方面，中国电信已经打造出了基于 AI 和大数据的无线网络智慧运营平台。

2.4 未来趋势

从设备控制的粒度上看，白盒化网络设备的发展至今经历了两个

阶段。在第一阶段，网络设备及其软件由网络所有者进行集中式控制。网络设备功能/协议可以被远程修改、配置。在此阶段，网络设备/软件/接口较为封闭，协议互通性差，转发逻辑固化，新协议/功能开发时间长，研发成本高，无法满足灵活多样的新型网络功能需求。因此，网络设备逐渐发展为设备架构开放、数据包转发可控的第二阶段，原有的固定 pipeline 转变为灵活可编程的 PISA（Protocol-Independent Switch Architecture）架构，伴随着 OVS、SONiC、FBOSS、FRR（FRRouting）、ONOS 等开源网络软件的崛起，不透明且封闭的网络开始变得透明且开放。

随着网络规模不断扩大，业务种类的不断增多，网络管控难度不断提升，对网络设备的管理需要摒弃由专人管理与维护的方式，构建包括 5G 在内的端到端白盒化开放体系，实现端到端、自顶向下、完全由软件定义的可编程。因此，需要聚焦功能和性能的监管，由少量网络所有者对网络资源进行包粒度级别的细粒度管控，优化网络的可靠性和安全性。在网络数据面研发端到端的白盒化网络设备操作系统，构建端到端白盒网络，支撑高效业务调度与感知（包括高效调度、高精度感知、高可控维护），实现全场景覆盖。采用先进的软硬分离、灵活可编程、按需而变的开放网络架构，力求满足不同行业的差异化、定制化网络需求，加速网络与实体经济的深度融合。针对网络管理面，构建智能化网络管理闭环，网络管理者仅需在顶部描述管理行为，网络会自动根据行为进行分区、编译和运行，将网络资源（包括云、ISP、

5G 网络) 视为可编程的载体, 通过软件自动化的方式进行日常的校验与实时检查。

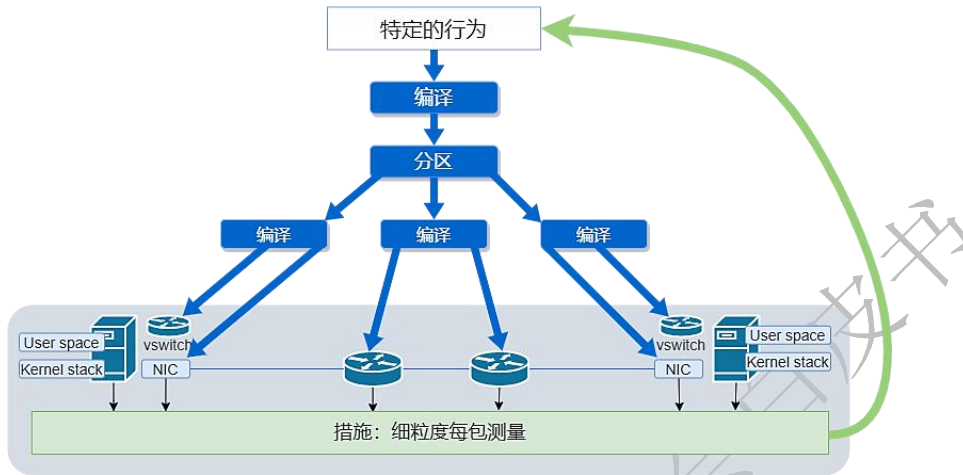


图 2-2 网络闭环控制

为了实现上述功能, 需要掌握以下三个方面的关键技术:

(1) **高可控维护**: 高性能 BFD (Bidirectional Forwarding Detection) 研究, 实现网络资源的毫秒级状态检测;

(2) **高精度网络感知**: 基于 INT (In-band Network Telemetry) [4]、Telemetry 等, 开展高精度网络测量研究, 实现带内网络遥测, 验证每个数据包或所有的状态是否“正确”;

(3) **高效网络调度**: 适用于大规模网络的 SR 路由机制, 实现流量带宽、路径高效调度与控制[5]。

如图 2-2 所示, 基于开放网络架构, 构建异构适配的、端到端白盒化网络数据面, 实现设备的无缝管理。在管理面, 网络拥有者的管控行为首先被分区编译, 交付给不同的网络设备执行。通过细粒度的测量技术, 获取执行反馈, 形成网络控制的闭环。这样一来, 网络能够根据网络规范进行实时管控与检测, 通过闭环控制对管控错误进

行更正，并且验证网络是否被正确构造。

第五届未来网络发展大会白皮书

三、 关键技术

白盒交换机涉及多个层次的相互配合，不仅包括硬件的选择与适配，还包括多项新型网络技术。为了梳理白盒交换机涉及的架构与技术，更好的推动该领域的技术与生态构建，本章将从软硬件解耦技术、可编程网络技术、硬件加速技术、白盒安全技术四个方面介绍白盒交换机设计的关键技术点，供读者共同交流和探讨。

3.1 软硬件解耦技术

传统网络设备的软硬件开发均由设备厂商提供，这种黑盒架构致使系统完全封闭，存在低利用率、高成本、低敏捷性和业务上线周期长等问题，难以适应未来网络的发展需求[6]。新兴的白盒设备打破了传统硬件设备系统的封闭性，将设备硬件与软件分离，能够将标准化的硬件配置与不同的软件协议进行混合并匹配，实现网络设备的白盒化，用户可以由此获得更加开放、灵活和更具成本效益地替代传统专有集成网络设备的方案。这就是软硬件解耦技术，是构建电信基础设施的新方法。

对于白盒设备的软硬件解耦需要在通用硬件上加载标准化接口、操作系统、控制面协议、管理面协议以及数据平面协议软件等等。硬件和软件必须来自可靠的供应商，硬件和软件的接口必须是开放且明确的。这些开放的接口是实现软硬件解耦的关键，而接口的安全和可信是新的生态系统足够强大、稳定和自我维持的保障。为实现开放、

解耦的目标和理想，众多的网络设备厂商、研究院、运营商和互联网用户在硬件和软件层面建立了完整的开放联盟和生态，让交换机从传统封闭架构走向开放架构，从而使得采用小型白盒硬件搭建大规模智能网络成为可能。

如图 3-1 所示，AT&T 将白盒交换机的生态系统可分解为四层：

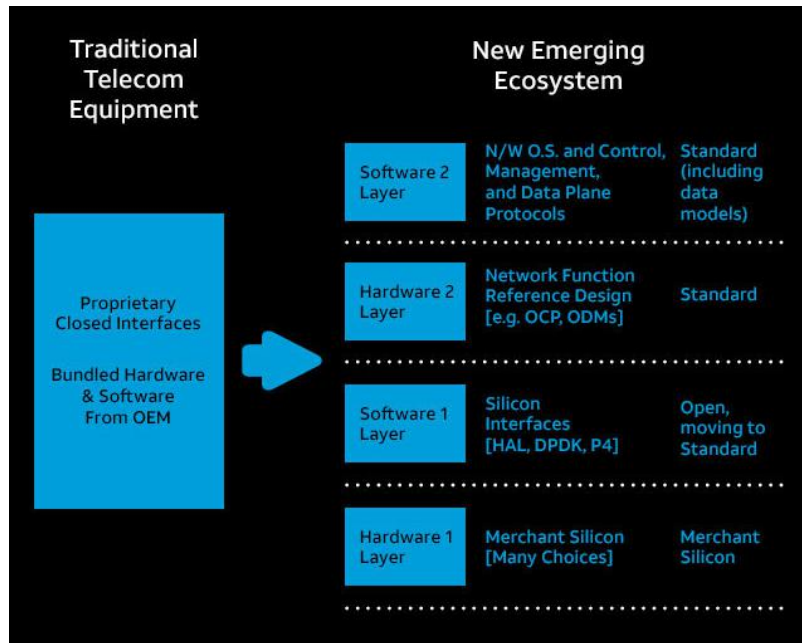


图 3-1 AT&T 白盒交换机生态系统¹

硬件 1 层 (Hardware 1 Layer)：商用芯片层，负责底层的交换转发。目前该层还没有硬性的标准。

软件 1 层 (Software 1 Layer)：芯片接口层，提取芯片的功能并向上提供服务。该层原则上需要标准化，但还需要一个过程。

硬件 2 层 (Hardware 2 Layer)：网络功能参考设计层，提供硬件设备网络功能设计参考。该层主要包括 OCP 项目制定的硬件设备网络功能的参考性设计。

¹ 该图来源于 https://about.att.com/innovationblog/scaling_white_box

软件 2 层 (Software 2 Layer): 网络操作系统和协议层, 负责实现控制和管理平面的功能。该层主要包括网络操作系统和上层的网络协议应用, 是最重要的一层。

这种新的细化的生态系统将传统一体化的软硬件解耦, 为服务提供商提供了几种不同的解决方案: 服务提供商可以对图 3-1 中右侧所示的各种功能作选择, 基本上拥有或控制了整个解决方案的集成、交付和支持。服务提供商可以拥有硬件解决方案, 并集成不同的软件模块。在此模式中, 服务提供商能够控制硬件以及软件的集成和维护。服务提供商的第二个选择是控制硬件和软件模块的设计和规格, 并使用第三方集成商进行制造、破解、集成和维护。第三个选择是确定要使用什么元素以及需要哪些功能, 并从单个实体购买完整的解决方案, 指定所需的元素和功能。

3.2 网络可编程技术

网络可编程是 SDN 的一个重要特性, 通过对网络进行编程, 可以灵活定义网络策略和行为, 便捷地实现任何新业务和新功能, 真正做到网络的软件定义。一般而言, 网络功能的实现需要控制面和转发面两部分, 白盒交换机需要在这两方面具备可编程能力。控制面较早的实现了集中式的可编程功能; 而大多数设备的转发面使用不可编程的黑盒转发芯片, 无法实现协议/功能的集中控制和定制。因此转发面的可编程技术是设备白盒化发展的关键。

3.2.1 控制面可编程能力

控制面主要对底层网络交换设备进行集中管理，包括状态监测、转发决策以及处理和调度数据平面的流量，对应图 3-1 则体现为 AT&T 四层结构中的软件 2 层（网络操作系统和协议层）。它向下通过南向接口协议对接硬件 1 层（商用芯片层），实现链路发现、拓扑管理、策略制定、表项下发等功能；向上则通过北向接口为上层业务应用以及资源管理系统提供灵活的网络资源抽象，开放多个层次的可编程能力。

一个典型的控制器体系架构如图 3-2 所示。南向接口层提供对各类南向接口协议（如 Open Flow、Netconf、OVSDB 等）的支持，控制器通过接口层的通道实现对硬件 1 层的管理；抽象逻辑层实现服务的抽象，完成多种通信协议的适配，为各模块和应用提供一致的服务；基础网络层既包括控制器内部的实现逻辑，如拓扑管理、链路计算等，也包括一些底层的网络实现逻辑，如 BGP（Border Gateway Protocol）Vxlan 等，是软件 2 层的核心所在；内置应用层提供如 L2、L3 网络，Overlay APP，服务链 APP 等基础的功能包；北向接口层实现 Restful API 的接口或者嵌入式的 API 供上层应用调用；配置管理层提供控制器服务管理，集群管理和图形化界面，实现在 web 界面中调用控制器的北向 API 对控制器进行配置。

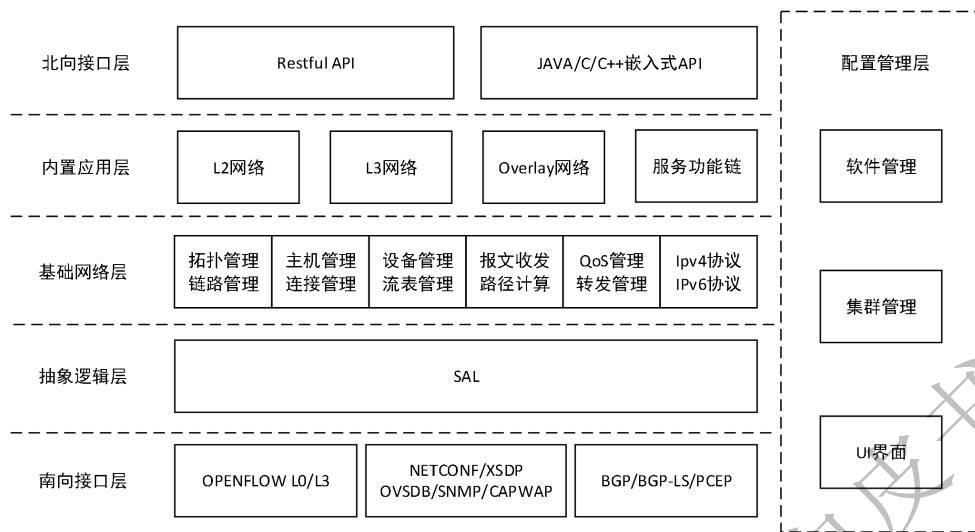


图 3-2 控制器体系架构

控制面可编程技术的发展将带来以下几点优势：（1）白盒交换机使用与服务器相似的网络操作系统，能够使用现有的服务器管理工具实现网络自动化，支持对开源服务器软件包的轻松访问，实现在交换机上使用与服务器完全相同的配置管理接口，提高创新速度；（2）将传统交换机特殊的网络环境变为较为通用的环境，从而对网络服务进行高效地拓展和管理，提高了白盒交换机的可编程性和对网络的可见性；（3）能够通过 API 和控制器，在交换机的网络操作系统中实现动态可编程，编写需要的网络功能（如网络分流器），从而在每一个交换机上减少硬件部署，做到对网络进行集中管理和监控。

3.2.2 数据面可编程能力

传统数据面将网络的报文处理和转发逻辑全部固化在硬件芯片中，由完全线速的芯片逻辑完成，从而使网络性能得到大幅提升，但却无法满足当今上层业务与控制软件对底层网络逐渐增加的特殊需

求。转发平面很大程度上受制于功能固定的 ASIC 芯片。如图 3-3 所示，固化的报文处理和转发逻辑使得上层业务系统只能在芯片能力所限定的范围内设计、开发，无法按需自行调整。同时，这类特殊需求通常具有复杂性与多样性，难以固化到芯片当中。因此可以利用可编程交换芯片实现如 NAT（Network Address Translation）或 SLB（Server Load Balancer）等常用的网络功能，这不仅降低业务部署复杂度，而且在提升网络处理性能的同时还确保了网络功能的可调整性与网络的灵活性。

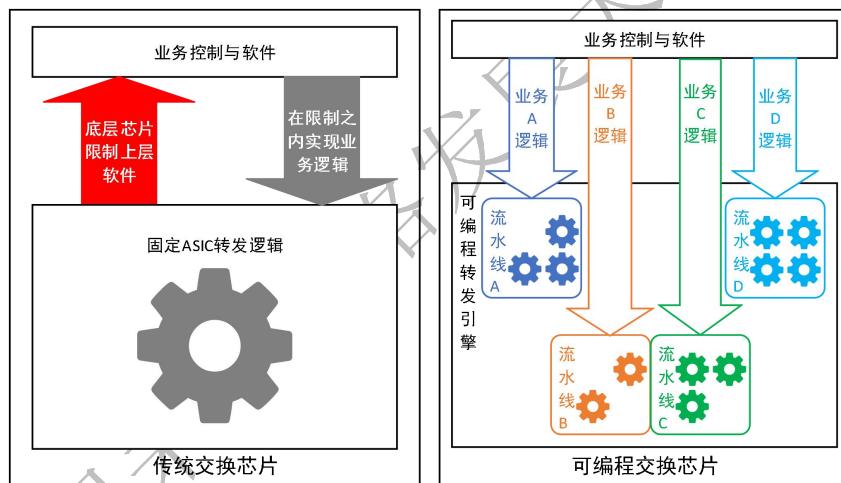


图 3-3 传统交换芯片与可编程交换芯片

可编程网络技术的核心是具备可编程特性的交换芯片，即芯片的报文处理和转发逻辑能够通过软件来按需调整。目前，可编程交换芯片的硬件载体是 ASIC 与 FPGA（Field Programmable Gate Array）的结合体。集成在芯片中的硬件资源（片载内存、流水线等）有限，在当前的网络应用环境中，虽然不能将所有的网络功能全部下移到芯片中实现，但将一些常见的网络功能以可编程的形式转移到芯片中的策

略已经得到了广泛应用。

可编程网络技术还需要高级编程语言的支持，使交换系统的设计开发者能按照上层业务系统的需求编写报文处理和转发逻辑。芯片将按照定制的逻辑流程来处理网络流量和报文。目前常见的交换机数据面编程语言有 P4、NPL（Network Programming Language）等。

在 2014 年，McKeown 教授等通过一系列论文，发布了一款全新的用于数据平面编程的高级编程语言——P4 语言[7]。P4 是一种专用的编程语言，其目标为协议无关性、目标无关性以及现场可重配置能力。它能够解决 OpenFlow 编程能力不足以及其设计本身所带来的可拓展性差的难题。首先 P4 定义数据包的处理流程，然后利用编译器在不受限于具体协议的交换机或网卡上生成具体的配置，从而实现用 P4 表达的数据包处理逻辑。

高性能可编程 ASIC 芯片与 P4 高级语言的出现，让网络拥有者可以自上而下地定义数据包的完整处理流程。这种数据平面可编程特性有助于快速的迭代开发，甚至可以直接通过打补丁的方式修复现有产品中发现的数据平面程序漏洞；它也可以帮助网络拥有者实现最适合其自身需求的具体网络行为；它还能使网络芯片供应商专注于设计并改进那些可重用的数据包处理架构和基本模块，而不必纠结于特定协议里错综复杂的细节和异常行为。

截至目前，P4 语言分为 P4₁₄ 和 P4₁₆ 两个大的版本。其中，2017 年 5 月 P4 社区 (p4.org) 发布的 P4₁₆ 语言是当前 P4 语言的最新版本。

随着 P4 交换机概念的兴起，以及 Barefoot 的 Tofino 系列可编程芯片所获得的高关注度，在 2019 年 6 月，Broadcom 正式发布了自有的 NPL 语言，并从 Trident4 芯片开始原生支持。

NPL 是一种数据平面高级编程语言，旨在提供一组适当的构造来描述数据平面数据包处理行为[8]。这些构造块的范围覆盖从允许指定单个控制信号的数据类型到允许将接口连接到复杂硬件的高级构造块。同时，NPL 语言没有固有地绑定到任何特定的硬件体系结构。它打算在多种硬件平台上实现，例如可编程 ASIC 芯片、智能网卡 SmartNIC（Smart Network Interface Card，智能网卡）、FPGA 以及纯软件交换机。但是，由于该语言的主要要求是构造与灵活硬件设备之间的映射，因此语言本身的整个可用结构集往往只集中在灵活硬件中对应的功能上，必须使用这些构造来完成应用程序开发。但目前 NPL 语言前景尚不明朗，对一些想要尝试可编程芯片的科研院所而言，成熟的 P4 语言更有吸引力。

3.3 硬件加速技术

在大部分的场景中，交换机负责处理网络数据包的传输，数据包最后抵达目的服务器后才进行相应的处理和计算。但是随着网络流量的急剧增长，限于 CPU 与交换芯片的性能瓶颈，现有的数据平面架构（如图 3-4 所示）已经不能很好满足用户对于低延迟、高传输的需求。

为解决上述问题，可在数据面上整合智能网卡、FPGA 等硬件加速卡，利用硬件加速技术实现网络流量卸载，降低整体网络延迟、减少 CPU/交换芯片的资源消耗，能显著提高网络的整体性能和服务质量。

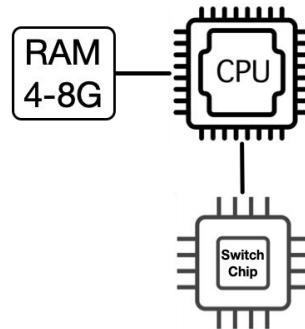


图 3-4 传统交换机架构

(1) 基于智能网卡的智能卸载

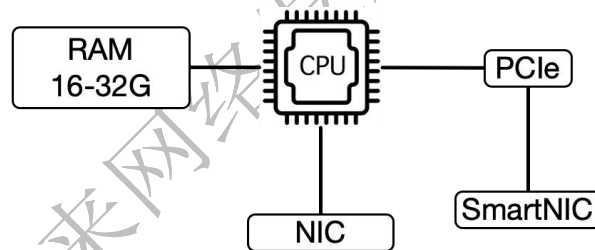


图 3-5 服务器+智能网卡型架构

如图 3-5 所示，数据面可采用 CPU+SmartNIC 的异构组合。其中，CPU 通过高速的 PCIe(Peripheral Component Interconnect Express) 接口与 SmartNIC 相连。在转发过程中，针对数据包需要特殊处理的部分（CPU 资源消耗巨大或硬件处理增益较大的网络功能）可以直接卸载至智能网卡[9]。这种组合方式不仅能实现正常的网络数据包转发，还强化设备的处理能力，可以有效提高白盒交换机的性能并减小网络延迟。

(2) 基于可编程加速卡和多 CPU 的硬件加速技术

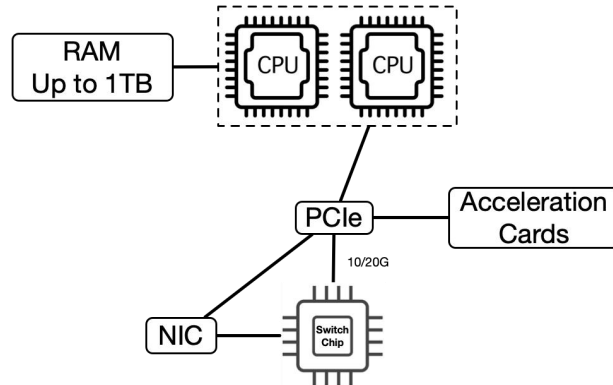


图 3-6 服务器+交换机+加速卡型架构

如图 3-6 所示，数据面可搭载多颗 CPU，实现 CPU 间的负载均衡。同时，通过 PCIe 总线可接入多个可编程硬件加速卡（包括 NIC，FPGA 等）。相较于前文与智能网卡的简单组合，这种组合方式可以连入更多计算、存储和网络资源。通过合理利用硬件资源（例如，NIC 可处理通用的网络数据包，需要特别处理的数据包可经由 FPGA 等加速卡处理），网络传输延时可以被极大的降低，网络传输质量可以得到较大的提升[10]。

3.4 白盒安全技术

白盒交换机的开放架构存在不可忽略的安全问题，例如针对 ONIE 的漏洞利用。ONIE 允许用户在不更换硬件的情况下，部署或更换网络操作系统（包括启动和恢复 Big Switch Networks、Cumulus Networks 等供应商的网络操作系统）。利用 ONIE 的漏洞和缺陷（包括缺乏身份验证和加密），攻击者可以在交换机的启动阶段（即操作系统完全加载之前），插入恶意代码。而操作系统的安全软件无法在

启动阶段运行，已加载的恶意代码会被认为是已知/良好的组件。即使检测到攻击，用户通过更换固件来移除恶意代码的成本也会很高。

针对 ONIE 的安全问题，目前的常规做法是添加关键功能控制潜在的安全问题。例如，Hellfire Security 公司已开始开发增强版 ONIE。新的版本将会增加身份验证和密钥强化来防止相关漏洞的利用。

除了白盒交换机开放架构带来的安全问题，数据面信息的加密传输也是网络安全考量之一。这一点主要体现在数据在硬件层传输过程是否安全。目前，博通等芯片制造商生产的芯片，性能卓越、应用广泛，可以有效保障数据传输的稳定和安全。大部分白盒交换机生产商（如思科、惠普、Arista、戴尔、IBM）都是采用博通芯片，用户可暂时不必担心数据传输的安全问题。

四、设备架构

白盒交换机分硬件和软件两个部分，其典型架构如图 4-1 所示。硬件一般包括交换芯片、CPU 芯片、网卡、存储以及外围硬件设备等，其接口、结构等需要符合 OCP 标准化规范；软件主要是指网络操作系统（Network Operating System, NOS）及其搭载的网络应用。在白盒交换机中，NOS 一般通过基础软件平台（如 ONIE）的引导完成安装，芯片接口层（如 SAI 等）则将交换芯片的硬件功能封装为统一的接口，解耦上层应用与底层硬件。具体而言，上层应用通过调用芯片接口定制底层转发逻辑，提供网络的可编程功能。本章将对各层主要技术做出介绍。

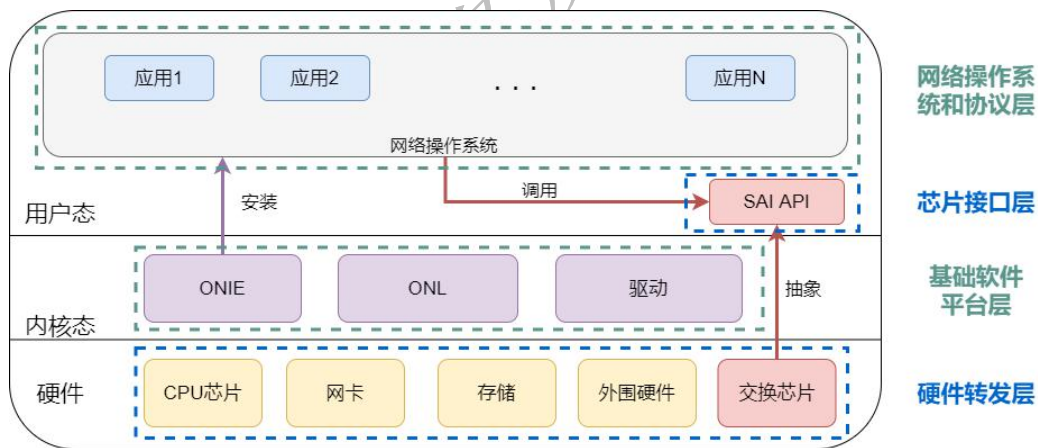


图 4-1 白盒交换机架构

4.1 硬件转发层

硬件转发层通常包含以下几类器件：（1）交换芯片：用于转发数据；（2）CPU 芯片：主要管控系统运作；（3）网卡：提供 CPU 侧管理功能；（4）存储器件：包括内存、硬盘等；（5）外围硬件：

包括风扇、电源等。其中，交换芯片负责交换机底层数据包的交换转发，是交换机最核心的部件。下面，本章对芯片厂商提供的可编程芯片做简要介绍。

4.1.1 Barefoot Tofino 系列芯片

Barefoot Tofino 交换芯片是业内首个支持协议无关交换架构（Protocol Independent Switch Architecture, PISA）的以太网交换 ASIC。其转发逻辑由加载到芯片上的 P4 程序决定[11]。图 4-2 为 PISA 架构处理流程图，数据包经由解析，通过多张流表进行匹配（match）和动作（action）操作，实现数据平面的协议无关转发。相对其它方式，PISA 提供的可编程性在实现过程中不会引入更多的功耗和成本。

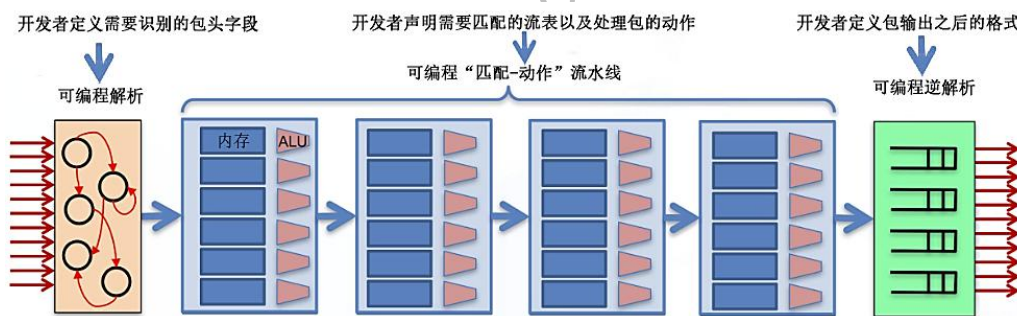


图 4-2 PISA 交换机架构²

Tofino 为 16nm 制程的芯片，采用 PISA 架构，支持 64x100G 端口，可实现高达 6.5 Tbps 的处理速度，使用 P4 可编程语言实现数据包转发平面的编程。后续 Barefoot 发布了 7nm 制程的 Tofino2 芯片，提供 32x400G 端口，并且总带宽提升到 12.8Tbs，更多比较如表 4-1

² 该图来源于 <https://events19.linuxfoundation.org/wp-content/uploads/2017/12/Tutorial-P4-and-P4Runtime-Technical-Introduction-and-Use-Cases-for-Service-Providers-Carmelo-Cascone-Open-Networking-Foundation.pdf>

所示：

表 4-1 Tofino 芯片比较

型号	Tofino	Tofino2
发行日期	Q2'16	Q4'18
光刻	16 nm	7 nm
端口	64x100G	32x400G
最大带宽	6.4 Tbps	12.8 Tbps
帧处理速度	4.8 B pps	6 B pps
共享包内存尺寸	22 MB	64 MB
CPU 接口	PCIe	PCIe
封装大小	60x60mm	71.5X66mm

Tofino2 芯片具有以下优点：1) I/O 灵活性：具备不同的串口与不同媒介；2) 拥有多个 SKU (Stock Keeping Unit)：芯片在子 SKU 不需要时可以将其移除；3) 单个芯片小，效率高；4) 电源效率：可在不需要时移除芯片。目前，基于 Tofino 芯片的白盒交换机主要用于学校、研究院等非营利性科研机构。国内也开始有互联网公司在尝试商用。

4.1.2 Broadcom 可编程系列芯片

Broadcom 的交换芯片分为 StrataXGS 和 StrataDNX 两个大类，前者包括 Trident、Tomahawk 系列，后者包括 Qumran、Jericho、FE3600/FE600 等，两者技术架构不同，StrataXGS 多用于盒式交换机，StrataDNX 多用于框式交换机，系列最新芯片比较如表 4-2 所示。

表 4-2 Broadcom 芯片比较

型号	Tomahawk4	Trident4	Jericho2c+
发行日期	4Q19	2Q19	4Q20
制程	7nm	7nm	7nm
端口配置	64 × 400GbE, 128 × 200GbE, 256 × 100GbE	32x 400GbE, 64x 200GbE, 128x 100GbE	144 x 50G
最大带宽	25.6 Tb/s	12.8 Tb/s	14.4 Tb/s
可编程性	fully-programmable	user-programmable	programmable

	streaming telemetry ³	forwarding ⁴	elements matrix ⁵
编程语言	嵌入式	NPL	C++

芯片的选型主要基于用户场景的区别，例如 StrataXGS 系列的芯片，Tomahawk 系列能够提供较大带宽，常用于汇聚交换机，适用于超大规模云网络，存储网络及 HPC 等场景；Trident 系列特性丰富，常用于 TOR 交换机，适用于数据中心、园区及无线交换等场景。StrtaDNX 系列是 StrataXGS 的补充，其中 Jericho2 常用于运营商以太网、边缘和核心路由器，适用于数据中心云以及企业园区场景。

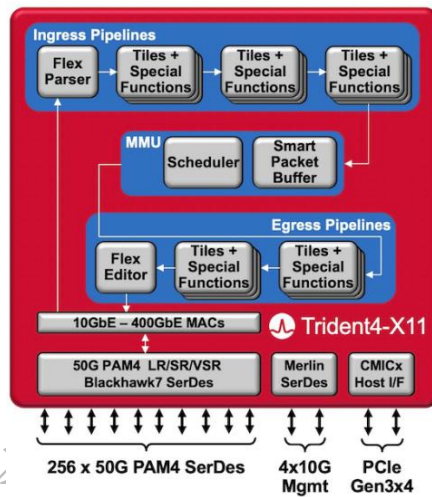


图 4-3 Trident4 芯片架构图

以数据面可编程芯片 Trident 4 为例，单片 7nm 完全可编程的芯片支持最大 12.8Tbps 带宽，其架构如图 4-3 所示。Trident 4 提供了独特的 runtime 可编程功能，支持在在运行期间对各种状态和功能进行更新，如接入控制策略表、遥测大数据的规模和类型，以及处理数据包丢失，都可以在不影响数据包流的情况下操作。Trident 4 能够将

³ 来源: <https://www.broadcom.com/products/ethernet-connectivity/switching/strataxgs/bcm56990-series>

⁴ 来源: <https://www.broadcom.com/products/ethernet-connectivity/switching/strataxgs/bcm56880-series>

⁵ 来源: <https://www.broadcom.com/products/ethernet-connectivity/switching/stratadnx/bcm88850>

多芯片系统整合到单一芯片上，从而大大降低系统的成本、功耗和软件复杂性，提高系统的可靠性。突破性的编译器编程架构为数据包转发和遥测提供了灵活性、可视性，同时保持了多样化的功能。

4.1.3 CISCO Silicon One 系列芯片

CISCO 的 Silicon One 系列芯片均采用相似的架构，支持 P4 语言进行转发面编程。用于路由的 Q200 芯片可达到 12.8Tbs 带宽，用于交换的 G100 芯片可达到 25.6Tbs 的带宽，其架构如图 4-4 所示。

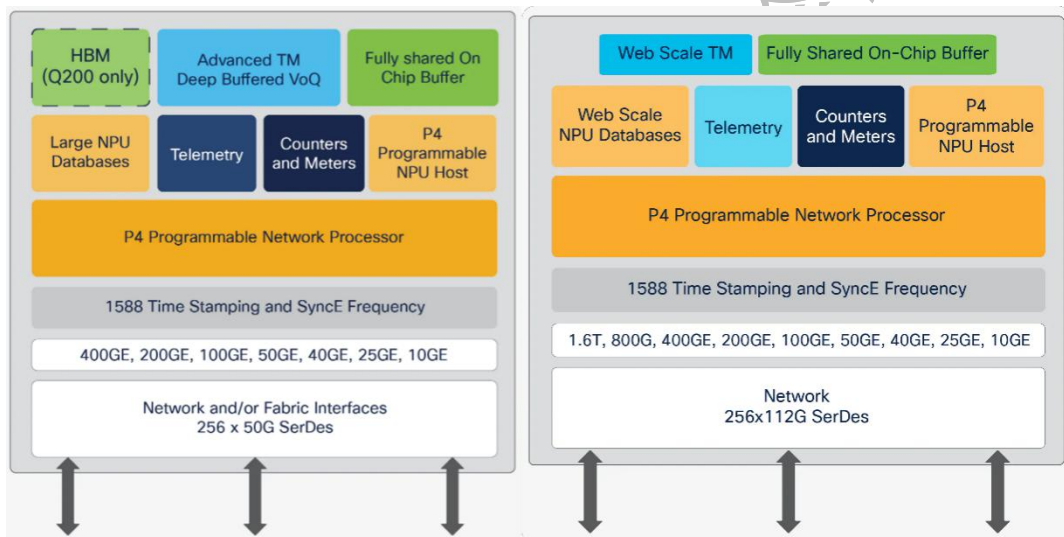


图 4-4 Silicon One Q200⁶/G100⁷芯片结构图

Silicon One 系列芯片可以大大简化网络基础架构部署，为整个网络基础架构中的所有应用程序提供一致的集成点；能够以线速处理数据包，满足服务提供商和 Web 规模提供商的路由和交换需求；还可以利用开源编程语言自定义功能，同时功能灵活性不会影响性能或电源效率。

⁶ 该图来源于 <https://www.cisco.com/c/en/us/solutions/collateral/silicon-one/datasheet-c78-744312.html>

⁷ 该图来源于 <https://www.cisco.com/c/en/us/solutions/collateral/silicon-one/datasheet-c78-744833.html>

采用基于 NP 芯片的 Silicon One 作为白盒交换机的核心，还有一个优势，即可以兼任交换芯片和交换矩阵芯片的角色。这大大简化了框式模块化交换机的设计制造等流程。目前 Silicon One 已经在 Facebook、Microsoft 等互联网公司有所应用，国内暂时没有厂商引入。

4.2 芯片接口层

芯片接口层旨在为上层应用提供标准化的芯片功能接口，实现芯片和上层应用的通信。为了使 NOS 与芯片接口平滑适配，芯片供应商提供的 SDK（Software Development Kit）需要与 NOS 集成，并提供标准接口来访问芯片。目前标准接口已被各种开源 NOS 广泛接受并广泛使用。由于它是标准的，因此开发人员只需要知道与供应商无关的标准 API 即可。目前主要使用的标准化接口有 OpenNSL（Open Network Switch Layer）/OF-DPA（Openflow Data Plane Abstraction）、SAI 等。

4.2.1 OpenNSL/OF-DPA

ASIC 的软件接口方面，OpenNSL 是面向传统 L2/L3 协议栈的 API，OF-DPA 是面向 OpenFlow 的 API。OpenNSL 将 SDK 映射到开放式北向接口，从而可以实现新应用的整合，并提供优化交换机硬件平台的能力。如图 4-5 所示，白盒设备本地的应用程序通过调用 OpenNSL API 来控制芯片的行为，也可以通过 JSON-RPC（JSON Remote Procedure Call）在远端进行调用。

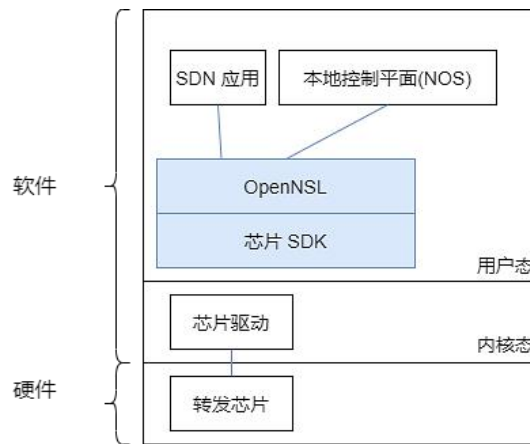


图 4-5 OpenNSL 架构

OF-DPA (openflow data plane abstraction) 是一个应用软件组件，其实现了 openflow 与 broadcom SDK 间适配层的功能。OpenFlow Agent 收到远端控制器的 OpenFlow 消息后，在交换机本地转化成 OF-DPA API 来控制芯片的行为，如图 4-6 所示。

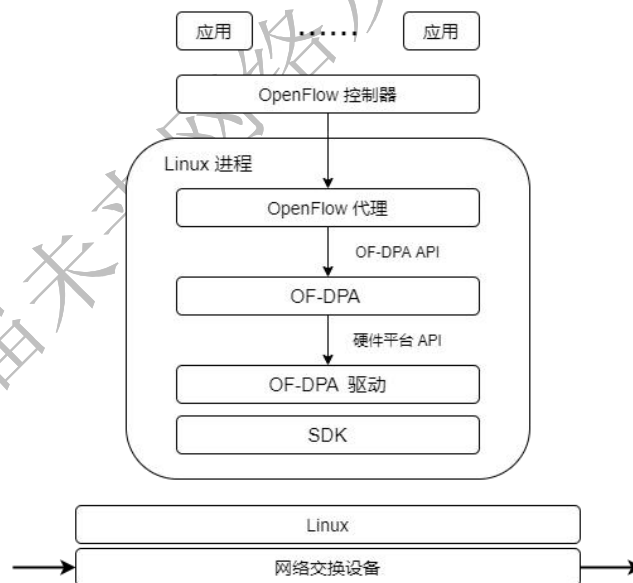


图 4-6 OpenFlow 到 OF-DPA 的转换

用户应用程序通过北向 API 从 openflow 控制器获取服务。控制器与其控制的每一个 openflow 交换机保持一个通道交互 openflow 协议信息。交换机上由 openflow 代理处理收到的 openflow 协议信息并

进行响应。OF-DPA 将所有与硬件匹配的表的状态都维护到 openflow，并向代理提供 OF-DPA API，使代理能根据 openflow 结构将 openflow 消息相对简单且无状态地转换为 OF-DPA API 调用。

OF-DPA 内部实际上依然使用 OpenNSL 库，但向上提供的 API 各不相同。虽然 OpenNSL 和 OF-DPA 分别在面向传统协议栈和 openflow 方面提供了丰富的 API，但白盒设备的能力最终仍受限于芯片提供的 API，而没有一个统一的接口标准。

4.2.2 SAI

SAI 是一种标准化的 API，涵盖多种功能，可以看作是一个用户级的驱动。使用者不需要担心硬件厂商的约束，不用关心其底层的交换芯片、网络处理单元或其是否是一个软件交换机，都可采用统一的方式进行管理适配。同时，其可以大大简化芯片厂商的 SDK。

在不同的 ASIC 芯片上，SAI 为上层应用提供了统一的 API 接口。SAI 的具体实现由不同 ASIC 芯片提供商负责，这其中以 Microsoft、Dell、Facebook、Broadcom、Intel、Mellanox 等为代表。使用者不需要关心网络硬件供应商的硬件体系结构的开发和革新，通过始终一致的编程接口就可以很容易地应用最新最好的硬件。

SAI 本质就是在各 ASIC 的 SDK（如上述提及的 OpenNSL）之上再做一层统一的抽象，芯片厂商研发的 ASIC 的 SDK 需要与这层抽象进行适配，使得转发应用能够在不同的 ASIC 上运行。SAI 向上为 NOS 提供了统一的 API 接口，向下可以对接不同的 ASIC。

如图 4-7 所示，SAI 未出现前，所有芯片都要通过自研 SDK 与上层的软件通信，相当于用“方言”与上层通信，因此需要白盒厂商自行适配不同的 ASIC。SAI 的出现将“方言”进行了标准化，由芯片厂商完成 ASIC 的适配。不同芯片用同样的语言与上层通信，使得白盒厂商能快速推出或迭代新产品。

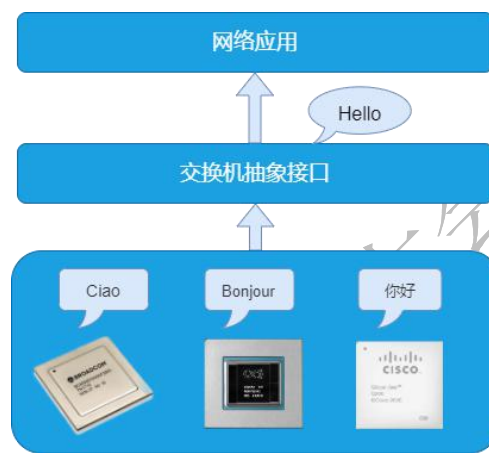


图 4-7 交换机抽象接口

SAI 的架构如图 4-8 所示。其中 ASIC SDK 被称为 Adaptor，一个 Adaptor 会根据 SAI 将 ASIC 中不同的功能适配到不同的 Adaptor Host 中。Adaptor Host 可以看作对 ASIC 中某个特定功能的抽象。例如，MAC 转发和 IP 转发是两个不同的 Adaptor Host。一个 ASIC 如果既支持 L2 又支持 L3，那么其 SDK 就需要把 L2 功能和 L3 功能分别适配到 MAC Adaptor Host 和 IP Adaptor Host 中。而不同 ASIC 的 L2 功能都会适配到 MAC Adaptor Host 中。SAI 对于转控分离的意义十分重大，目前已经有很多 ASIC 厂商提供了对于 SAI 的支持。

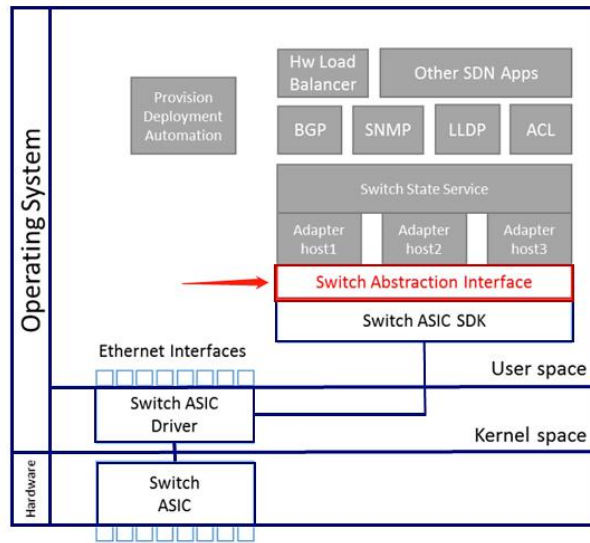


图 4-8 SAI 对不同的 ASIC SDK 进行抽象⁸

4.3 基础软件平台层

基础软件平台层主要连接了白盒设备的底层硬件和上层网络操作系统，提供基础的安装环境、驱动等软件，目前主要的规范是 OPC 提出的 ONIE 和 ONL。

4.3.1 ONIE

ONIE (Open Network Install Environment) 为白盒交换机提供了一个开放的安装环境，实现了交换机硬件和网络操作系统系统的解耦，支持在不同厂商的硬件上引导启动 NOS。当使用 ONIE 的网络设备首次启动时，引导加载程序会启动内核来运行 ODE (ONIE Discovery and Execution) 程序。ODE 可以通过如本地文件、DHCP、IPv6 邻居、mDNS/DNS-SD 等方式来定位和下载(通过 HTTP 或 FTP) 操作系统安装程序，一旦找到就会运行该程序。

在 ONIE 出现之前，交换机使用预装的专有操作系统，与特定的

⁸ 该图来源于 https://hackmd.io/@ldotrg/ryC_c6NtD

硬件强相关。这导致终端用户被绑定到设备生产商的垂直供应链上。用户不仅要为 NOS 支付高额的软件使用和维护费用，也很难针对自身的业务进行个性化定制、指定硬件形态。现在 ONIE 已经成为网络硬件行业的主流安装环境，OCP 为使用 ONIE 的交换机定义了一套硬件标准。终端用户不仅可以自由选择 NOS，而且可以根据标准自行设计白盒交换机。

针对安装 ONIE 的交换机的硬件配置，OCP 制定了一系列标准，如以下几点：

(1) 非易失性存储

每个 ONIE 系统必须包括非易失性存储，其中包含制造商分配的重要产品数据。存储介质可以采用 EEPROM（Electrically Erasable Programmable Read Only Memory），NOR-FLASH 扇区，NAND 闪存扇区等形式。EEPROM 包含诸如分配给系统的 MAC 地址、序列号、制造商的日期等信息。一个 EEPROM 模块的示例如表 4-3 所示：

表 4-3 EEPROM 模块示例

Field Name	Size in Bytes	Value
ID String	8	“TlvInfo”
Header Version	1	0x01
Total Length	2	Total number of bytes that follow
TLV 1	Varies	The data for TLV 1
TLV 2	Varies	The data for TLV 2
.....
TLV N	Varies	The data for TLV N
CRC-32 TLV	6	Type = 0xFE, Length = 4, Value = 4 byte CRC-32

其中各字段用处如下：1) ID String：用作格式的简单初步检查；2) Header Version：表示版本号；3) Total Length：表示数据总长度，

用于确定要读取的数据量和确定 CRC 的位置；4) TLV：由类型码字段、长度字段和值字段三个子字段组成，类型码定义了信息的类型和格式，如 0x21 表示产品名称，0x2A 表 MAC 地址数量，0xFE 表示 CRC-32 校验，长度表示值字段的字节数，值字段表示类型码对应的值。

(2) 串行控制台波特率

ONIE 项目要求交换机默认的串行控制台波特率是 115200，用户可以根据需要在使用时自由进行更改，出厂时设置为默认的 115200。

(3) 硬件面板和 FRU 编号

ONIE 项目为机器上的钣金丝印推荐以下约定：1) 交换机端口标记数字以“1”开始；2) 交换机端口数字从上到下，从左到右标记；3) 现场可更换单元（FRU）标记以“1”开始。

(4) 硬件文档

ONIE 项目建议交换机提供一页快速入门指南，包括以下内容：1) ONIE 版本号和认证号码/日期；2) 交换机前后视图的模板；3) 温度传感器的位置布局；4) 控制台端口引脚输出，波特率/设置信息；5) 通用的 ONIE 安装说明。

4.3.2 ONL

ONL 是 Big Swtich 的商业交换机 NOS——Switch Light OS 的开源版本。它本质上是一种 Linux 发行版，其在 Linux 的基础上添加了

大量与交换机硬件相关的驱动程序，如 USB、GPIO（General Purpose Input Output）、CPLD（Complex Programmable Logic Device）、Flash、PCI、serial、RTC（Real-Time Clock）、EEPROM、DMA（Direct Memory Access）等，可供白盒交换机硬件厂商共享相同的交换机硬件驱动，在一定程度上减少了开发工作量。

ONL 是建立在开放网络硬件上，向下一代模块化 NOS 提供基础操作系统，为交换硬件提供管理接口。它使用 ONIE 来安装到板载闪存中。如图 4-9 所示，标准 ONL 发行版中的组件包括：Debian Linux 内核、一组设备驱动程序、安装脚本和具有 net-boot 功能的零接触网络启动加载程序，并针对各种裸机交换设备定制了增强的网络启动功能。

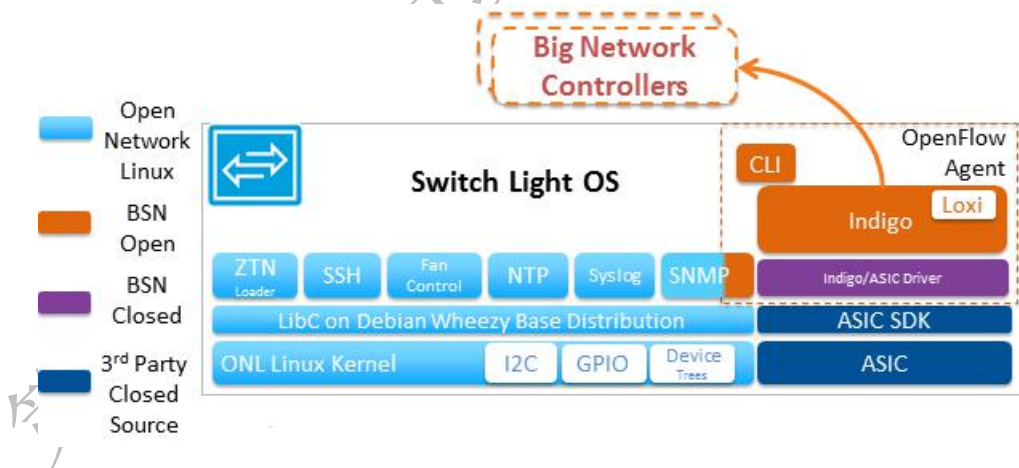


图 4-9 ONL 构成⁹

ONL 特点是简单，它只提供基本的 Linux 运行环境和最基本的网络功能，交换机软件厂商可以基于 ONL 开发复杂的网络功能，甚至基于 ONL 推出新的操作系统。对于希望在开放式网络交换机之上

⁹ 该图来源于 <https://screeninet.wordpress.com/2014/05/27/white-box-switches-overview-and-current-models/>

构建自己的网络操作系统(NOS)的最终用户，ONL 将充当基础组件的角色。它使硬件供应商能够更快地访问多个与平台无关的转发操作系统。

ONL 最初定位是用于白盒交换机，目前已在多个开放网络平台上得到验证，如建立在 ONL 之上的 Switch Light OS 已部署在数百个大型电信、企业和 SaaS 数据中心中。目前，ONL 也可以用于其它类型的网络设备，如白盒 OLT，未来将会有更多基于 ONL 的白盒网络设备出现。

4.4 网络操作系统和协议层

网络操作系统和协议层位于白盒交换机架构的最顶层，直接面向用户，主要包括网络操作系统和上层的网络协议应用。开源网络操作系统主要包括 SONiC、DANOS、Stratum、Switch Light OS 等；商用网络操作系统主要包括思科 IOS/NX-OS、Juniper JUNOS、H3C Comware 等。

4.4.1 SONiC

SONiC 是一个由微软牵头建立的开源项目。它将传统交换机操作系统软件分解成多个容器化组件。用户（包括运营商等）可根据自己的需求，在通用硬件上加载不同的功能组件（软件协议堆栈），实现系统的敏捷开发和无缝快速升级。

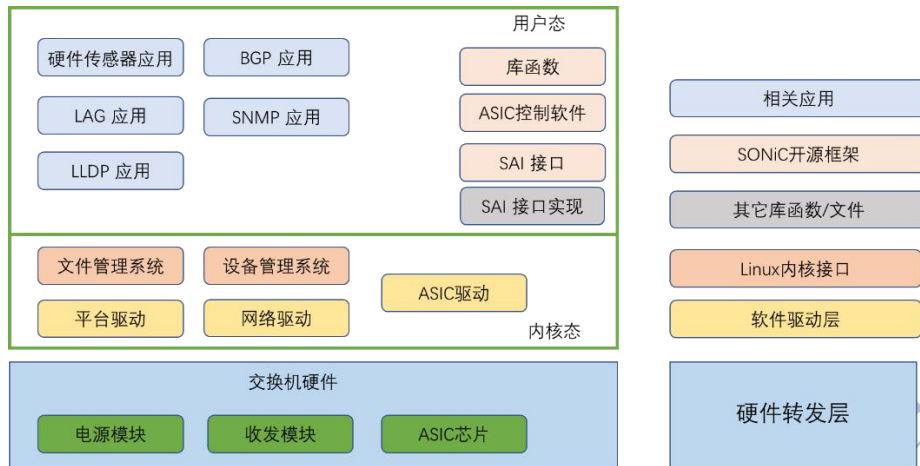


图 4-10 SONiC 系统架构¹⁰

如图 4-10 所示，SONiC 采用标准化的 SAI 接口实现软硬件的解耦。SONiC 大量使用了现有的开源项目和开源技术，如 Docker, Redis, Quagga 和 LLDP 以及自动化配置工具 Ansible、Puppet 和 Chef 等。SONiC 的开源特性使得用户可以借鉴相关项目经验（如微软在云数据中心网络的经验），开发出针对性的网络应用，减少开发流程，降低开发成本。

需要注意的是，SONiC 并不是一个完整的操作系统，而只是一个构建交换机网络功能的软件集架构，它需要运行在 Base OS—ONL 上。在部署过程中，需要将 SONiC 和 Base OS、SAI、ASIC 平台对应的驱动打包制作成为一个文件。这个文件才是可直接安装到白盒交换机的 NOS 镜像。SONiC 目前已经可以运行在 Accton/EdgeCore、Dell 和 Mellanox 等厂商的白盒交换机上。

4.4.2 UniNOS

UniNOS 由紫金山实验室研发，是国内首个兼容 **Broadcom** 等异

¹⁰ 该图来源于 https://github.com/Azure/SONiC/blob/master/images/sonic-architecture_new.png

构芯片的白盒化开放网络设备操作系统，能够支撑构建端到端白盒网络，实现全场景的异构网络覆盖。

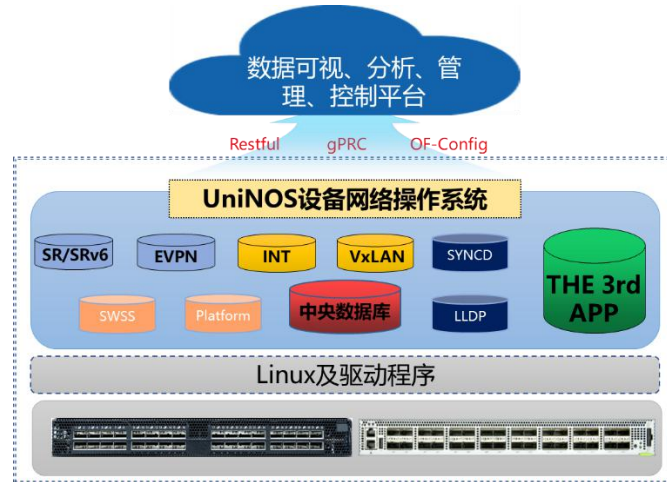


图 4-11 UniNOS 架构

如图 4-11 所示，UniNOS 采用软硬分离、灵活可编程、按需而变的开放网络架构，其特点包括（1）能实现软件与软件、软件与硬件的解耦；（2）易集成扩展第三方软件；（3）解耦遵循 SAI、RestFul 等标准化接口；（4）支持 P4、NPL 等芯片编程语言；（5）以数据库为核心构建的容器化架构。UniNOS 利用统一抽象 SAI 接口和数据面统一抽象模型，支持多芯片适配，能够兼容博通、Intel、Mellanox 等异构交换芯片，实现设备无缝管理。另外，UniNOS 具有骨干级业务调度与感知能力，可通过骨干级 SR 路由机制，实现流量带宽、路径高效调度与控制；使用 INT（Inband Network Telemetry）等技术，实现高精度网络测量；通过高性能 BFD，实现毫秒级链路状态检测。

在网络性能优化方面，UniNOS 使用 RDMA（Remote Direct Memory Access）替代传统的 TCP/IP，可实现零拷贝+内核旁路+硬件

卸载的高性能协议栈，提供超高吞吐、超低时延和零丢包的无损网络特性。其次，为构建灵活弹性的超大规模网络资源池，打破数据中心地域的限制，UniNOS 支持基于 SDN 集中式控制+BGP EVPN(Ethernet Virtual Private Network) 分布式控制。另外，通过结合新型骨干网可编程路由技术、大规模确定性网络技术，UniNOS 将白盒设备的部署场景拓展至大网级网络，并提供电信级可靠性保证（包括城域网、骨干网、5G 承载网等场景）。

4.4.3 DANOS

如前文所述，DANOS 的前身是 AT&T 的 dNOS。2017 年，dNOS 改名为 DANOS，并托管给 Linux 基金会。目前，DANOS 已得到了各种 Linux 基金会团体和成员的支持，包括 Broadcom，Inocybe，Metaswitch 和 Silicom。

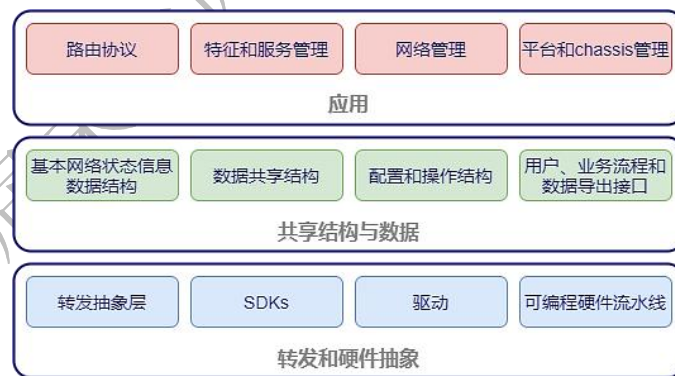


图 4-12 DNOS 功能层和组件¹¹

如图 4-12 所示，DANOS 由硬件和软件组件组成。在软件部分，DANOS 定义了一套标准的应用程序编程接口和框架，主要涉及基本

¹¹ 该图来源于 https://about.att.com/content/dam/innovationblogdocs/att-routing-nos-open-architecture_FIN_AL%20whitepaper.pdf

操作系统、控制和管理平面以及数据平面，分别对应五种组件（FRR, SONiC SAI, Open Switch, Stratum 和 P4）。

DANOS 的硬件部分至少包括一个通用 CPU，用于运行基本操作系统，控制和管理平面以及用例所需的任何软件数据平面。DANOS 要求 CPU 硬件必须支持虚拟化，软件组件可在虚拟环境中运行。除通用 CPU 之外，硬件部分还可以包括其它加速设备以支撑性能加速功能，例如 ASIC, NPU (Neural-Network Process Units), FPGA 等。

4.4.4 Stratum

Stratum 是谷歌于 2018 年 3 月推出的一个开源项目，其目的在于实现一个软件定义的数据平面，其目标是提供一个适用于通用白盒交换机的开放软件系统。

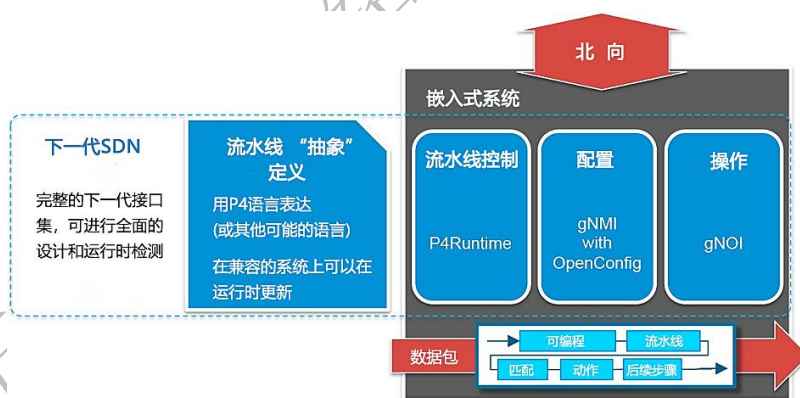


图 4-13 Stratum 系统架构¹²

如图 4-13 所示，Stratum 为白盒交换机构建一个开放的、最小化的、商用化的操作系统。为实现转发设备的可互换性和转发行为的可编程性，Stratum 集成了若干下一代 SDN 接口 (Next-Gen SDN

¹² 该图来源于

<https://www.opensourceforu.com/2018/03/open-source-sdn-switching-platform-google-support/>

interfaces), 包括 P4, P4Runtime, gNMI/OpenConfig 和 gNOI 等。因此, Stratum 简化了开源代码的复杂程度, 可实现完整生命周期的自动化网络管控。在可编程性方面, Stratum 可通过一个“契约”(用 P4 语言表达), 精确定义数据平面支持的转发行为。目前来看, Stratum 主要适用于 5G 移动通信和运营商边缘云平台以及云 SDN 数据平面。

4.4.5 DENT

DENT 是由 Amazon 发起的新 Linux 基金会项目, 其目的是为小校园网和零售商店提供简单的低成本网络操作系统。DENT 利用 Linux 内核、Switchdev 和其它 Linux 开源项目, 平等对待所有的底层硬件设备(包括用于网络和数据处理的 ASIC 和 CPU), 简化 APIs、驱动、接口等逻辑实体, 构建一个标准化的网络操作系统。

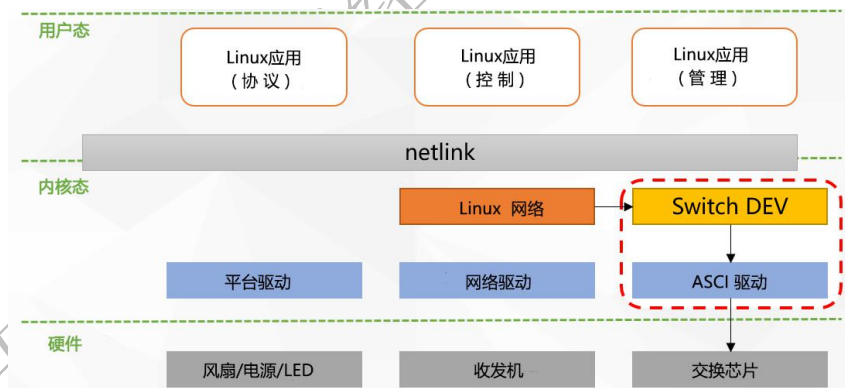


图 4-14 DENT 架构

如图 4-14 所示, 以 spectrum driver 卸载 fib 表为例说明。(1) sp 内核 driver 注册 fib notify: register_fib_notifier, 回调函数为 mlxsw_sp_router_fib_event; (2) kernel 中添加一条 fib 表处理函数 fib_table_insert 中, 通过 call_fib_entry_notifiers 最终调用上面回调函

数。整个过程较为简单，与服务器系统相似，只需要芯片厂商提供内核驱动，利用 Linux 内核进行软硬件屏蔽，应用程序无需感知硬件。Switch DEV 接口目前已经支持基本的 fdb、neighbour、fib、vxlan、vlan、接口、acl、组播等常见网络功能。

4.4.6 总结与对比

表 4-4 NOS 总结与对比

NOS	编程语言	目标应用场景	主要参与者	硬件抽象接口
SONiC	swss 和 syncd : C++ mgmt: Go、Python、C protocol: C	数据中心	Azure, barefoot, Broadcom, Mellanox, arista, Centec, dell, Alibaba, tencent	SAI
UniNOS	Go...	数据中心、城域网、骨干网、5G	紫金山实验室、北京邮电大学	uSAI
DANOS	C、Perl、Go、Python	VNF、接入网	AT&T	FAL
Stratum	C++	数据中心	google, Tencent, 中国联通, cisco, VMware, barefoot, broadcom, mellanox	P4runtime
DENT	未开源	校园网、边缘网、零售商店	Amazon, Cumulus, Delta, Marvell, Mellanox, WN C, edge-core	switch DEV

如表 4-4 所示，NOS 的开源项目几乎都由 OTT 或运营商主导，聚焦于实现软硬解耦，简化管理流程，像管理服务器一样管理交换机。项目主导者一般从事 IT 行业，习惯性地项目中引入数据库发布订阅、消息中间件、RPC 等机制。然而可以看出，不同项目的硬件抽象接口各异，缺乏统一标准，如 DENT 项目完全使用处于内核态的交换芯片驱动接口，实现过程严重依赖 linux 内核。

五、 应用场景

随着白盒交换机的硬件、NOS、控制管理软件和数据面可编程技术逐步成熟，白盒交换机逐渐在多个场景得到应用，如数据中心、新型承载网、5G 云等。

5.1 数据中心场景

数据中心通常使用高性能以太网交换机，实现计算资源之间的快速通信。大型数据中心的网络需求体现在三个方面：高度稳定性、高度可管可控、高性能低成本。数据中心场景需求明确，对软件特性的要求相对简单，拓扑结构明确，非常适合采用白盒交换机。通过部署白盒交换机，相关厂商/企业能够自主掌控软硬件运营体系，搭建更低成本、更为可靠、高度可控、高度自动化、智能化的数据中心网络，最终提供更优质服务。从国内以及全球数据中心网络设备市场来看，网络软件化和硬件白盒化已经成为了趋势[12]。

在数据中心场景中，构成底层网络的交换机在不同的位置上扮演不同的角色，图 5-1 所示是一个典型的 Spine-Leaf 架构的数据中心网络，在这个网络中，所有的交换机被分为 Spine 交换机、服务器 Leaf 交换机（Server ToR）、网关 Leaf 交换机（Gateway ToR）三种不同的角色。

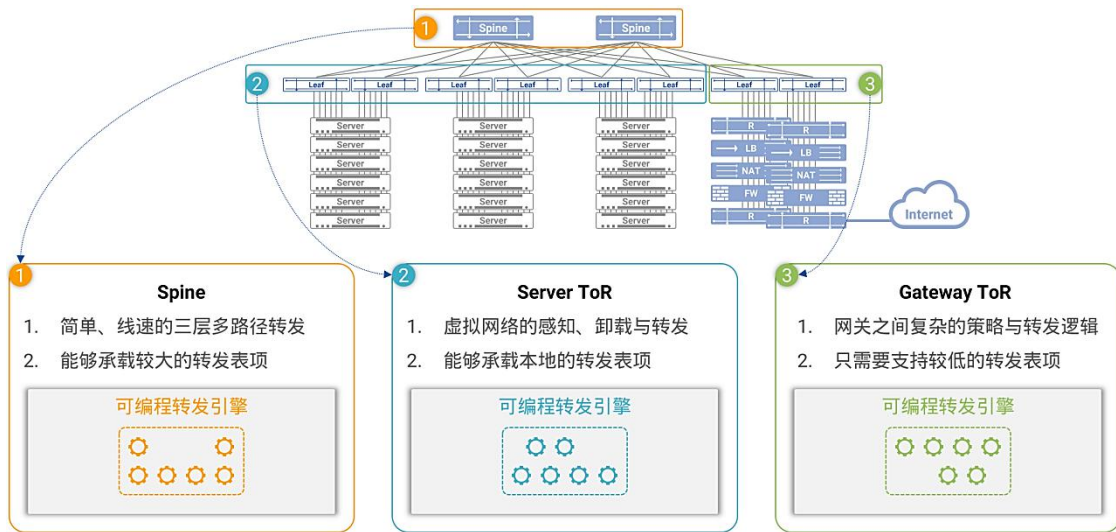


图 5-1 Spine-Leaf 架构

Spine 交换机：主要承载所有 Leaf 交换机之间的所有流量，在尽可能多地将系统资源划分给 FIB 表的同时，需要具备大容量、多租户的区隔能力，即要支持创建和维护相当数量的 BD（Bridging Domain）和 VRF（Virtual Route Forwarding）。

服务器 Leaf 交换机：主要承载虚拟计算节点之间的东西向流量，完成各种内部业务或应用的高速、大容量交换。在此过程中，需要尽可能多地为 FIB 表分配资源，以承载尽可能多的二层 MAC 转发表项和三层主机转发表项。

网关 Leaf 交换机：主要承载对外业务的虚拟计算节点与外部世界（互联网或企业内联网）之间的南北向流量，是业务对外提供服务的转发通道，需要尽可能多地将资源划分给三层路由表，同时能够支持各种 NSF（如 NAT、DDoS 防御等）。

上述的三种角色在数据中心网络的不同位置承担了不同任务，其需要处理的流量的特征、所需要花费硬件资源的配比等都是不同的。

配备可编程芯片的白盒交换机能够通过加载不同的功能组件，承担不同角色的不同转发逻辑，实现硬件资源不同的分配。不仅能够将一些高价值的 NSF 直接承载在底层网络上，并以线速处理完成；而且能够根据交换机所承担角色的不同，进行有针对性的优化，以更好的性能为上层业务提供支持。

5.2 新型承载网场景

新型承载网是面向未来新型网络业务，基于 SR-MPLS (Multi-Protocol Label Switching) /SRv6、EVPN、网络切片、确定性网络等技术，对固网和移网业务进行综合承载的新型城域网络，其架构如图 5-2 所示[13]。

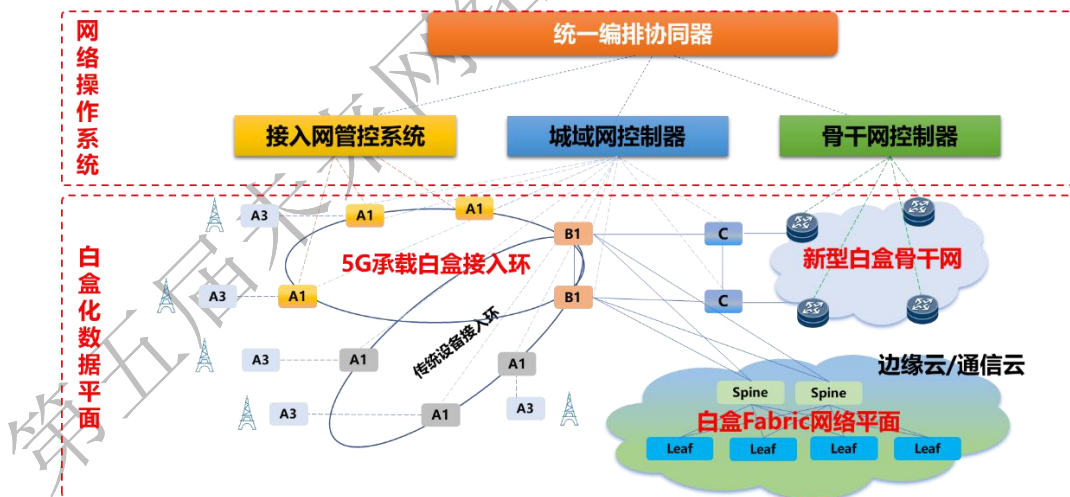


图 5-2 白盒化 IP 承载网网络架构

新型承载网的主要业务场景共包括两类，第一类是 5G 移动回传业务，实现将基站流量接入到核心网网元；第二类为承载网宽带业务，包括家庭宽带、IPTV (Internet Protocol Television)、互联网专线、VoIP

(Voice over Internet Protocol) 等业务等。与现有的网络不同，新型承载网强调网络协议与网络功能的可编程、确定性、定制化、高性能需求，能提供软硬切片和 QoS (Quality of Service) 保证的专线服务。目前，白盒交换设备正在弱化交换机和路由器的设备形态，从接入承载网络到新型骨干网络，白盒交换设备可以做到全场景全兼容。因此，各大运营商开始陆续尝试基于白盒交换设备部署新型承载网。

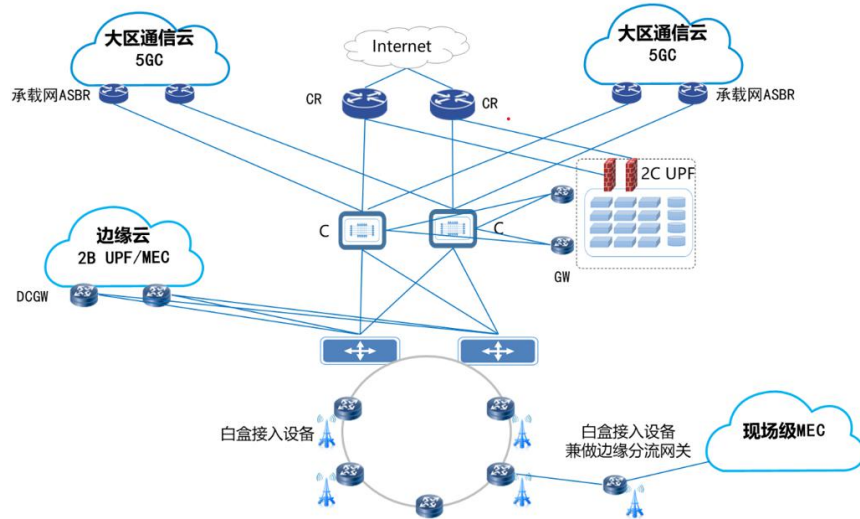


图 5-3 5G 移动回传和专线业务场景

5G 移动回传业务涉及的网络架构如图 5-3 所示，白盒设备主要可应用在接入部分和 UPF 部分。移动回传业务通过 MPLS/BGP L3VPN (Virtual Private Network) over SR/SRv6 隧道进行承载，部分需要专线承载的业务通过 EVPN L2VPN over SR/SRv6 隧道进行传输。在保证电信级的网络可靠性方面，运用网络保护/恢复技术，在部署过程中要求设备支持主控冗余，支持 NSR (Non-Stop Routing) 功能，VPN/EVPN FRR、Ti-LFA、SR-TE HSB 等功能。同时，要求支持 BFD、RFC2544、RFC1564 等层次化的 OAM 机制，实现故障的

检测和性能的检测。为满足业务确定性、高性能要求，设备需要支持 10GE、50GE、100GE 端口以及 FlexE（Flex Ethernet）子端口，具备 SyncE、1588v2 等时钟同步、时间同步功能。综上，针对 5G 回传业务，设备需具备 NSR、硬切片、高精度时钟同步等新型网络的功能。

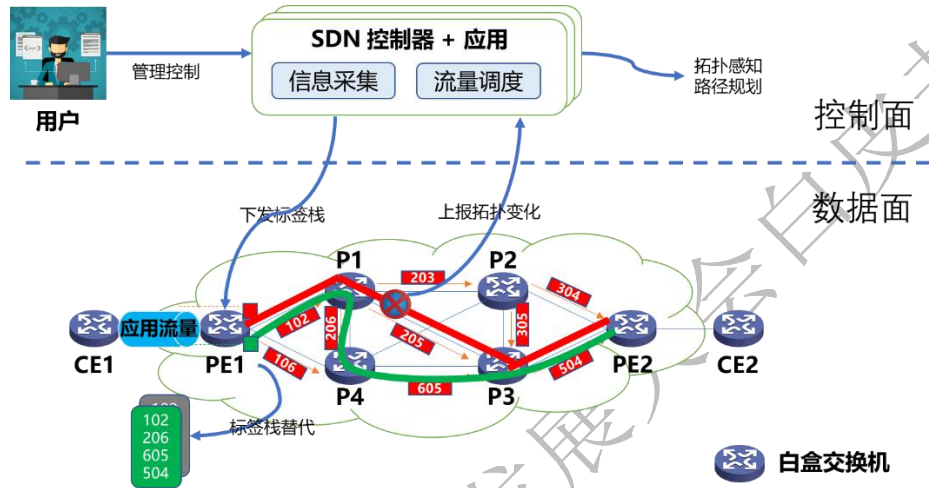


图 5-4 承载网宽带业务场景

承载网宽带业务涉及的网络架构如图 5-4 所示，主要包括承载家宽、互联网专线、IDC、语音、移动互联网的固网承载部分。其中，白盒设备主要可应用在骨干网传输部分，并采用控制面和数据面解耦的结构。在控制面中，网络应用以 NFV 的形式部署于省或地市通信云中。在数据面中，为满足固网宽带高可靠、高带宽、确定性等业务需求，新型承载网引入 SR-MPLS、SR-POLICY、IGP (Interior Gateway Protocols) for SR、BGP-LS 等多种新型协议，支持无缝 BFD 路径快速检测、故障快速保护倒换、全网路径探测和性能测量等多种特性，实现电信级可靠性、确定性网络传输质量保障、智能运维、集中式算路引擎等功能。

5.3 5G 云场景

5G 云是指通过云计算和网络功能虚拟化等技术利用虚拟化网络单元构建云化的 5G 通信网络，同数据中心一样，其一般具有 Spine-Leaf 结构，只不过容器化的是网络设备。白盒交换机采用开放的网络架构，可以灵活的适配、管理虚拟化的网络设备，提升 5G 通信网络整体的服务质量和网络质量。按照云化 5G 的发展现状，5G 云主要包含 5G 通信云和 5G 边缘云两个应用分支[14]。

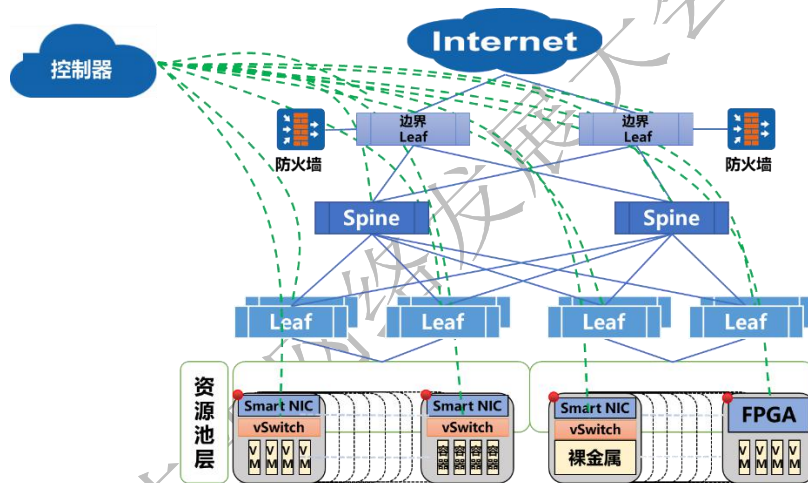


图 5-5 5G 通信云白盒化场景

5G 通信云是指提供端到端通信服务的基础 5G 云。如图 5-5 所示为 5G 通信云场景的白盒交换机解决方案，控制器可以采用 SDN 方式实现底层网络的自动部署和上层网络的定制化编排；边界 Leaf 和 Spine 可以部署白盒交换机，采用可编程转发芯片实现软硬一体化高性能网关设备，基于 RDMA 打造高吞吐、低延时、零丢包的无损网络，利用新型流量探测技术如带内网络遥测（INT）实现网络流量的可视可查；资源池层可以利用 SmartNIC 等新型可编程网卡，实现

OVS 流量的硬件卸载加速，提升转发性能并降低转发时延，同时也可以利用 FPGA 的可编程能力从计算、存储、网络全方位为业务定制化服务进行加速。

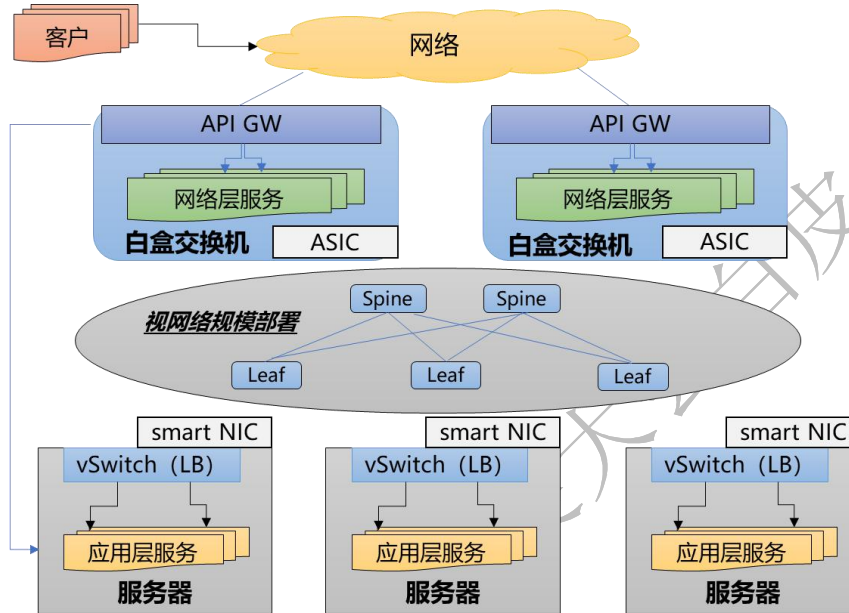


图 5-6 5G 边缘云白盒化场景

5G 边缘云是指在近用户侧进行通信处理的 5G 云。如图 5-6 所示为 5G 边缘云场景的白盒交换机解决方案，白盒交换机可以为交换机内部的网络层服务以及服务器内部的应用层服务统一提供网关服务，而网关又可以充分结合交换机和服务器的功能特点提供代理、聚合、鉴权、限流、日志、服务链跟踪、服务注册与发现、监控、熔断、降级、负载均衡等多种融合服务，减少不同设备的重复功能，提供一致的服务接口，优化网络部署和运维结构，提升网络可靠性、经济性和简约性。同时基于白盒交换机 ASIC 芯片和服务器 SmartNIC 可以实现 L4 负载均衡的两级硬件卸载，显著提升服务器负载均衡的性能。

六、典型案例

6.1 数据中心案例

传统数据中心通常使用高性能以太网交换机来实现计算资源之间的快速通信。对于底层网络设备，大型数据中心的需求体现在三个方面：高度稳定性、高度可管可控、高性能、低成本。白盒交换机采用成熟开放的软硬件架构，能够经济高效的部署运维，可以充分满足数据中心场景的需求。

6.1.1 解决方案

H3C 应某互联网客户数据中心组网运营自动化的需求，定制开发多款基于 Broadcom Trident3/Tomhawk3 的白盒交换机。于 2020 年，进行了实际部署工作，其架构如图 6-1 所示。

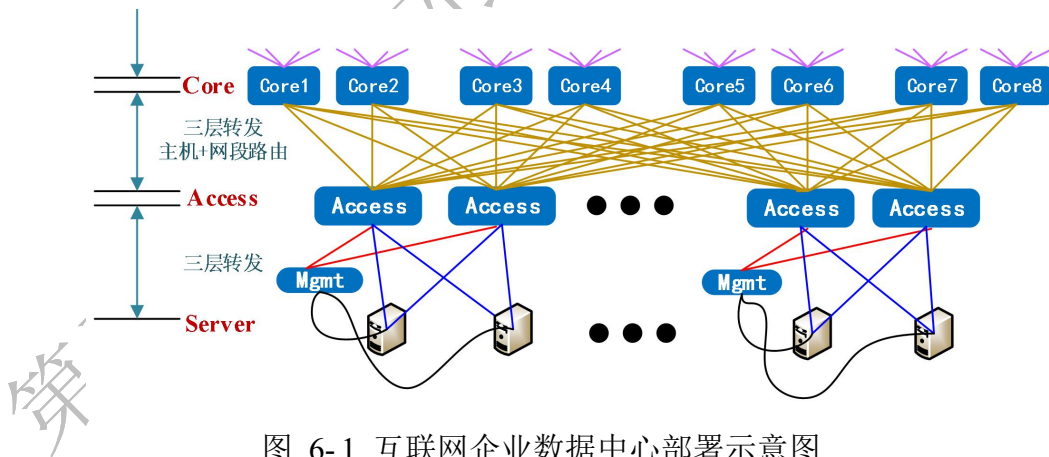


图 6-1 互联网企业数据中心部署示意图

在部署过程中，数据中心网络分为 3 层架构，外来业务流从核心层（Core）汇入，通过接入层（Access），访问应用（server）。设备方面，接入层白盒交换机采用 Trident3 转发芯片，为接入层提供 48x25G+8x100G 接口；核心层交换机采用 Tomahawk3 转发芯片，提

供 128x100G 转发能力；核心层设备的上连则采用商用框式交换机。软件层面，白盒交换机的 NOS 由客户基于 SONiC 自行开发。组网方面，白盒交换机之间使用 BGP 协议打通，进行全三层转发。

6.1.2 部署效果

目前，通过部署白盒交换机来实现数据中心的可编程化，客户得到了如下几方面的性能增益：（1）统一自动化管理平台；（2）BUG 修复和特性更新迭代更快，升级流程自动化；（3）业务升级更为灵活便捷；（4）故障的快速自愈。在使用白盒交换机后，数据中心的自主可控能力得到了极大的增强，并且更进一步催生了关于网络自主运营的新概念和新想法。未来依托白盒交换机设备，客户将要进一步优化架构仿真、故障预测以及动态验收等能力，从而将设备稳定性的保障由实施环节提前到设计环节，最终构建出完全自动化、高度可靠的数据中心网络。

6.2 校园网机房案例

校园网是为学校师生提供教学、科研和综合信息服务的宽带多媒体网络。首先，校园网是一个具有交互功能，且专业性很强的局域网络，需为学校教学、科研提供先进的信息化教学环境。其次，校园网应承载多种业务，包括教务、行政和总务管理功能。综上，校园网具有高带宽、设备异构、信息结构复杂、可靠性和安全性要求高等特点 [15]。

6.2.1 解决方案

针对校园网的相关要求，钰登科技于 2019 年联合 Big Switch 等公司针对美国某公立大学学校的校园网机房升级任务提出了解决方案，并定制了多款白盒交换机。项目的实际部署如图 6-2 所示。

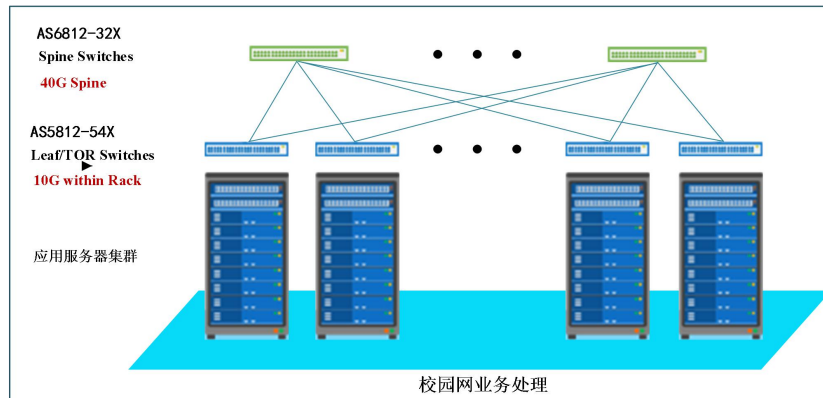


图 6-2 校园网机房部署示意图

在部署过程中，将校园网机房分为 2 层架构，外来业务统一经由 2 台 spine switch 汇入，通过 4 台 leaf switch 定点访问后台的应用服务器。在部署过程中，交换机的型号及参数如下所示：

(1) **Spine switch:** 2 台 AS6812-32X 40GbE 数据中心交换机，带 32 个 QSFP+40G 端口，作为 Spine 交换机，具有 40G 下行链路到 Leaf 交换机；

(2) **Leaf switch:** 4 台 AS5812-54X 10GbE 数据中心交换机，带有 48 个 10G SFP+端口和 6 个 40G QSFP+上行链路端口，用作分支交换机，具有到 Spine 交换机的上行链路和到控制器和计算节点的下行链路。

6.2.2 部署效果

基于白盒化交换机的部署方案除了能提供无缝、安全的连接以及管理各种 IT 应用程序和基础架构外，还支持未来网络规模的无缝升级，以便容纳更加庞大的学院科系及众多的学生的需求。同时，白盒交换机还支持通过集中供应整个班级的沙盒应用程序来降低运维人员的操作复杂性，减少运维开销。

6.3 政务网机房案例

电子政务网是政府组织处理电子政务的网络。它可以通过计算机、网络和通信等现代信息技术手段，实现政府组织机构和工作流程的优化重组，以便全方位地向社会提供优质、规范、透明、符合国际水准的管理与服务。政务网具有内容广泛、内容层级复杂、可靠性和安全性要求高等特点。

6.3.1 解决方案

2019 年，针对美国某城市的政务网升级，钰登科技联合 PICA8 等企业，定制了多款白盒交换机，提出了定制化的部署方案，建立了一个白盒化政务系统。该系统能够覆盖 9 个城市建筑物（包括两个数据中心），支持该市的所有运营信息系统，包括电话、公共安全系统和应急运营中心等，部署方案如图 6-3 所示。

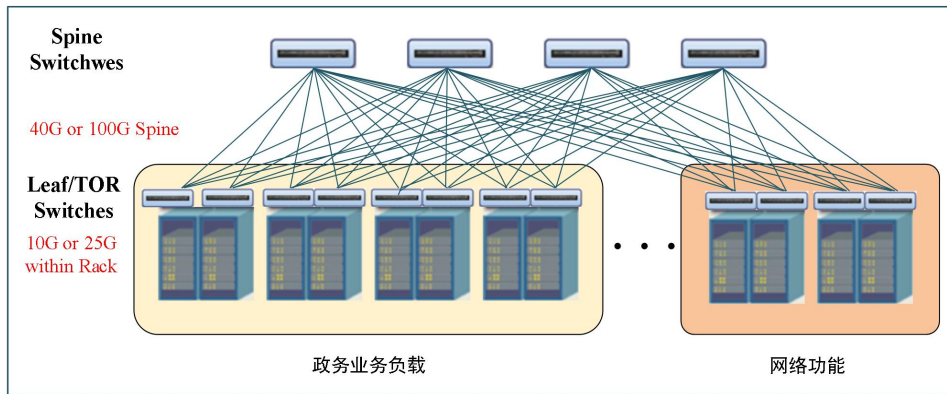


图 6-3 政务网机房案例部署示意图

在部署过程中，各个政务站点分为 2 层架构，包括约 50 台的 Edgcore Networks 交换机，支持到每个配线柜的两个 10G 至 20G 连接，每个配线柜的总容量为 40G。外来业务统一经由 Spine Switch 汇入，通过 4 台 Leaf Switch，访问后台应用服务器。服务器之间通过逻辑隔离的方式，提供可靠的访问、接入功能。部署交换机的型号及参数如下：

(1) **AS5812-54X (10G 数据中心交换机)**：可在 48 x 10GbE 端口和 6 x 40GbE 上行链路提供全线速二层或三层转发。该交换机可以作为架顶式交换机部署，也可以作为 10GbE 或 40GbE 分布式主干的一部分进行部署，支持基于硬件的 VXLAN 和 NVGRE 隧道功能，能够形成无阻塞的折叠式 CLOS 数据中心网络，是虚拟化数据中心的理想架顶式交换机。

(2) **AS4610-54T (1G 数据中心交换机)**：具有 48 个 10/100/1000BASE-T 端口，4 个 10G SFP+ 上行链路端口和 2 个 20G QSFP+ 堆叠端口（堆叠能力取决于 NOS），完全支持 IEEE 802.3at 和 IEEE

802.3af 以太网供电(PoE),包括无风扇设计,冗余的可热插拔 AC PSU 和端口到电源的气流,是数据中心机架式交换机的理想选择。

6.3.2 部署效果

目前,通过部署白盒交换机,借助扁平的两层 Leaf-Spine 架构,系统管理器能够获得更大的控制权,同时支持时延和性能的进一步优化和提升。与专有解决方案相比,使用白盒化部署方案可节省 35% 至 40% 的成本,且与其余的传统网络基础架构完全向后兼容。市中心新网络解决方案的成功部署也使政府决定逐步向市外扩展,在未来朝更多的城市与乡镇发展。

第五届未来网络发展大会白皮书

七、挑战与建议

今天，全球的电信运营商们都坚信可编程技术（包括 SDN 和白盒交换机）将深刻地改变通信网络领域的市场格局。白盒交换机采用开放的设备架构和软硬解耦的思想，提升了设备可编程性、灵活性，可以有效支撑未来新型业务对网络可定制、高性能、可编程、确定性的需求，在我国网络战略发展中具有重要地位[1]。目前，从国内以及全球数据中心网络设备市场来看，网络软件化和硬件白盒化已成为趋势，运营商承载网、5G 云白盒化已成为未来运营商网络的重要演进趋势。从产业生态上看，白盒设备已初步形成覆盖芯片、ODM、硬件、软件的完整骨干白盒网络生态，从商用可编程芯片到白盒硬件设备的标准化，从统一芯片接口到开源的交换操作系统，白盒生态的发展已经完全具备产业化能力。

7.1 挑战

值得注意的是，白盒交换机产业技术的发展也面临着一些挑战，具体如下。

白盒交换机网络创新和实践探索不成熟不广泛。目前，在利用白盒交换机搭建高性能、可定制、确定性的新型网络方面，业界已经积累了一定的实践经验。运营商一般通过 OpenFlow 等南向协议对白盒化的数据平面设备（如交换机等）进行集中控制。然而 OpenFlow 是协议相关的，仅能兼容几种常规的网络协议，无法灵活地适配用户所

需的新协议与新功能[16]。同时，白盒交换机的部署实践主要集中在云数据中心场景。针对大规模网络场景或大网级场景，如新型城域网、5G 通信云，白盒化交换机的技术研究和产业探索还处在初级阶段。

白盒交换机关键技术研究及开源开放不全面不深入。当前，白盒交换机的研发工作主要聚焦于软硬件解耦、网络可编程等方面。在软硬件解耦方面，业界一般在 OCP 等开源组织框架下，制定相关的软硬件技术规范。然而，目前相关的标准（例如 OCP 发布的标准）仅规定了最基本的硬件要求，缺乏更加具体明确的硬件细节。相关接口或组件，如 SAI 接口等，只提供二进制可执行文件，其代码细节仍待开源。在可编程技术方面，设备研发严重依赖芯片专用的编程语言。数据平面缺少通用的编程语言、模型和接口。另外，在异构加速及内生安全方面，投入精力相对较少，尚未形成完善的技术体系和有效的产业化发展。

白盒交换机市场规模及占用率仍有待提升。目前，白盒交换机的年出货量随着各大运营商的认可而快速增长。但相对传统的一体化交换设备，白盒交换机的市场总体占比并不高，且主要服务于规模较大的云服务商和电信运营商。规模较小或者技术较差的中小企业限于认知或成本因素，依然选择使用传统的一体化交换机。另一方面，即使是云服务商和电信运营商，受限于现网环境（传统交换机耦合性太高，更新难度大，功能需求不明确等），通常将白盒交换机和传统交换机混杂部署，无法形成一个完整的、端到端的白盒化数据平面。这种部

署方式极大的限制了白盒交换机的性能优势（如网络的可编程性、灵活性、确定性等），也阻碍了业界对白盒交换机功能和性能的挖掘和探索。

7.2 建议

为了推动白盒交换机的技术落地，促进网络技术的发展和变革，本文针对白盒交换机的发展提出以下建议。

推进白盒交换机硬件及接口标准化建设，明确一体化方案设计。面对未来多样化的业务需求，建议构建端到端、全覆盖的白盒化网络数据平面。为此，需要保证不同配置（硬件/软件版本）的白盒交换机能够互通及互操作。目前 OCP 仅规定了最基本的硬件要求，无法满足异构白盒交换机互通所需的基础硬件标准和接口标准。因此，更加详细具体的硬件标准化工作迫在眉睫[17]。同时，相关接口（如 SAI）的设计也需跟随新型网络功能而进行不断更新与演进。上述两方面都亟需各大开源组织、上下游企业统筹规划和推动。

推进白盒设备应用与融入，创新构建网络新型架构。需要考虑设计面向全场景、全覆盖的白盒化开放式网络架构，探索白盒与骨干网融合的网络新范式，支持 5G、4K/8K、AR/VR 等新业务发展，提供低成本、高效率、软硬分离、灵活可编程、按需而变的网络架构[18]。同时，加快新一代信息网络技术开发和自主标准的推广应用，支持适应下一代网络架构的转发芯片与 NOS 的研制和应用，带动新型网络设备产业和新兴信息服务及其商业模式的创新发展。

针对白盒设备关键技术及应用瓶颈进行攻关与突破。重点突破软硬件解耦技术、可编程网络技术、硬件加速技术、白盒安全技术等关键技术，支撑多样化网络需求，抢占技术创新高地，引领网络产业形态变革[19]。同时，需要构造智能化网络，突破高性能、可扩展、服务化的网络操作系统核心技术。通过开源社区的建设，构建完备的白盒生态和技术体系，支撑白盒设备成为新一代网络架构中的基础设施，构建开放的网络生态环境。

加强白盒设备应用成果输出，赋能多产业应用场景。白盒设备作为未来网络中的基础设施，将对全球网络构建产生重大影响[20]。因此，需要加强面向数据中心、新型承载网、5G云等重大场景的示范应用成果输出，开展核心技术在这些重大场景的一系列示范应用。通过加强技术和示范应用的发展和推广，整合我国电信运营商、设备厂商、互联网公司、研究机构及高校的资源优势，培养更多优秀的技术型人才，形成自主可控的白盒产业生态链，赋能国家新基建和新产业发展。

附录 A：术语与缩略语

英文缩写	英文全拼	中文释义
5G	5th Generation Mobile Communication Technology	第五代移动通信技术
ASIC	Application Specific Integrated Circuit	专用集成电路
BD	Bridging Domain	桥接域
BFD	Bidirectional Forwarding Detection	双向转发检测
BGP	Border Gateway Protocol	边界网关协议
CORD	Central Office Rearchitected as a Datacenter	电信端局重构为数据中心
CPLD	Complex Programmable Logic Device	编程逻辑器件
DANOS	Disaggregated Network Operating System	分解式网络操作系统
DMA	Direct Memory Access	直接内存访问
EEPROM	Electrically Erasable Programmable Read Only Memory	电气可擦除可编程只读存储器
EVPN	Ethernet Virtual Private Network	以太网虚拟专用网
FBOSS	Facebook Open Switching System	Facebook 开放交换系统
FlexE	Flex Ethernet	灵活以太网技术
FPGA	Field Programmable Gate Array	可编程门阵列
FRR	FRRouting	一种路由协议包
GPIO	General Purpose Input Output	通用型之输入输出
HPC	High Performance Computing	高性能计算
IGP	Interior Gateway Protocols	内部网关协议
INT	In-band Network Telemetry	带内网络遥测
IPTV	Internet Protocol Television	网路协议电视
JSON-RPC	JSON Remote Procedure Call	JSON 远程过程调用
MPLS	Multi-Protocol LabelSwitching	多协议标签交换
NAT	Network Address Translation	网络地址转换
NFV	Network Functions Virtualization	网络功能虚拟化
NG-SDN	Next Generation SDN	下一代 SDN
NOS	Network Operating System	网络操作系统
NPL	Network Programming Language	网络编程语言
NPU	Neural-Network Process Units	嵌入式神经网络处理器
NSR	Non-Stop Routing	不间断路由
OAM	Operation Administration and Maintenance	运行管理与维护
OCF	Open Compute Project	开放计算项目
ODCC	Open Data Center Comittee	开源数据中心委员会
ODE	ONIE Discovery and Execution	ONIE 发现与执行
ODL	OpenDaylight	开放网络控制器

OF-DPA	Openflow Data Plane Abstraction	Openflow 数据平面抽象
ONAP	Open Network Automation Platform	开放网络自动化平台
ONF	Open Networking Foundation	开放网络基金会
ONIE	Open Network Install Environment	开放网络安装环境
ONL	Open Network Linux	开放网络 Linux
ONOS	Open Network Operating System	开放网络操作系统
OpenNSL	Open Network Switch Layer	开放网络交换层
OVN	Open Virtual Network	开放虚拟网络
OVS	OpenVSwitch	开放虚拟交换机
P4	Programming Protocol-independent Packet Processors	协议无关的可编程包处理器
PCIe	Peripheral Component Interconnect Express	快速外设组件互连标准
PISA	Protocol-Independent Switch Architecture	协议无关交换机架构
QoS	Quality of Service	服务质量
RDMA	Remote Direct Memory Access	远程直接数据存取
RTC	Real-Time Clock	实时时钟芯片
SAI	Switch Abstraction Interface	交换机抽象接口
SDK	Software Development Kit	软件开发工具包
SDN	Software Defined Network	软件定义网络
SD-WAN	Software-Defined WAN	软件定义广域网
SKU	Stock Keeping Unit	库存量单位
SLB	Server Load Balancer	服务器负载均衡
SmartNIC	Smart Network Interface Card	智能网卡
SONiC	Software for Open Networking in the Cloud	云开放网络软件
SR	Segment Routing	分段路由
TIP	Telecom Infra Project	电信基础设施项目
VoIP	Voice over Internet Protocol	基于网络协议传输的语音
VPN	Virtual Private Network	虚拟专用网络
VRF	Virtual Route Forwarding	虚拟路由转发
WAN	Wide Area Network	广域网

参考文献

- [1] 黄韬, 霍如, 刘江, 等. 未来网络发展趋势与展望[J]. 中国科学: 信息科学, 2019, 8(49): 941-948.
- [2] 刘曦. 数据中心网络 SONiC 白盒技术的发展趋势[J]. 通信世界, 2019, No.826(33):44-45.
- [3] 闫耀华. 云数据中心中虚拟交换机 OpenvSwitch 的性能优化技术研究及实现 [D]. 北京邮电大学, 2017.
- [4] Tan L, Su W, Zhang W, et al. In-band network telemetry: A survey[J]. Computer Networks, 2021, 186: 107763.
- [5] Ventre P L, Salsano S, Polverini M, et al. Segment routing: A comprehensive survey of research activities, standardization efforts and implementation results[J]. IEEE Communications Surveys & Tutorials, 2020.
- [6] Bera S, Misra S, Vasilakos A V. Software-defined networking for internet of things: A survey[J]. IEEE Internet of Things Journal, 2017, 4(6): 1994-2008.
- [7] Bosshart P, Daly D, Gibb G, et al. P4: Programming protocol-independent packet processors[J]. ACM SIGCOMM Computer Communication Review, 2014, 44(3): 87-95.
- [8] He B, Dong L, Xu T, et al. Research on network programming language and policy conflicts for SDN[J]. Concurrency and Computation: Practice and Experience, 2017, 29(19): e4218.
- [9] Wang S, Meng Z, Sun C, et al. SmartChain: Enabling high-performance service chain partition between SmartNIC and CPU[C]//ICC 2020-2020 IEEE International Conference on Communications (ICC). IEEE, 2020: 1-7.
- [10] Vajaranta M, Oinonen A, Hämäläinen T D, et al. Feasibility of FPGA accelerated IPsec on cloud[J]. Microprocessors and Microsystems, 2019, 71: 102861.
- [11] Singh S K, Rothenberg C E, Patra G, et al. Offloading virtual evolved packet gateway user plane functions to a programmable ASIC[C]//Proceedings of the 1st ACM CoNEXT Workshop on Emerging in-Network Computing Paradigms. 2019: 9-14.
- [12] Bari M F, Boutaba R, Esteves R, et al. Data center network virtualization: A survey[J]. IEEE communications surveys & tutorials, 2012, 15(2): 909-928.
- [13] McCauley J, Liu Z, Panda A, et al. Recursive SDN for carrier networks[J]. ACM SIGCOMM Computer Communication Review, 2016, 46(4): 1-7.

- [14]Condoluci M, Mahmoodi T. Softwarization and virtualization in 5G mobile networks: Benefits, trends and challenges[J]. Computer Networks, 2018, 146: 65-84.
- [15]Nguyen V G, Kim Y H. SDN-based enterprise and campus networks: a case of VLAN management[J]. Journal of Information Processing Systems, 2016, 12(3): 511-524.
- [16]左青云, 陈鸣, 赵广松,等. 基于 OpenFlow 的 SDN 技术研究[J]. 软件学报, 2013(5):1078-1097.
- [17]梅雅鑫, 舒文琼. 网络"黑""白"之争[J]. 通信世界, 2018, No.787(29):16-17.
- [18]解云鹏, 王江龙, 雷波. 5G 时代 数据中心网络演进正当时[J]. 通信世界, 2019, No.826(33):34-35.
- [19]黄志兰,樊勇兵,陈楠,何晓武.overlay SDN 实现异构兼容的关键技术[J].电信科学,2016,32(11):112-118.
- [20]徐雷, 侯乐. ICT 全面融合 白盒开启运营商网络转型之路[J]. 通信世界, 2018, No.787(29):22-23.

第五届未来网络发展大会