

电信级开放网络设备关键技术研究 与研发实践

Key Technology Research and Development Practice of Carrier-grade Open Network Devices

徐博华, 庞冉, 张学茹, 王泽林, 曹畅(中国联通研究院, 北京 100048)

Xu Bohua, Pang Ran, Zhang Xueru, Wang Zelin, Cao Chang(China Unicom Research Institute, Beijing 100048, China)

摘要:

网络设备的能力开放是SDN/NFV技术的自然延伸,其中白盒交换机已在云数据中心领域验证了其可行性和先进性。近年来随着5G网络的规模建设,RAN设备、CPE设备、承载网络路由器设备的架构和能力开放已成为新的研究热点。首先从开源社区、产业现状等方面介绍了开放网络设备的背景,并进一步分析了开放网络设备研发所涉及的关键技术,最后阐述了中国联通在5G承载网络接入路由器领域所开展的能力开放创新实践,并对后续研发方向进行了展望。

关键词:

5G承载;开放网络设备;网络操作系统;SDN
doi:10.12045/j.issn.1007-3043.2022.04.003
文章编号:1007-3043(2022)04-0010-06
中图分类号:TN915
文献标识码:A
开放科学(资源服务)标识码(OSID):



Abstract:

The capability opening of network devices is a natural extension of SDN/NFV technology, among which white box switch has proved its feasibility and advanced in the field of cloud data center. In recent years, with the large-scale construction of 5G network, architecture and capability opening of RAN equipment, CPE equipment and bearer network router equipment have become a new research hotspot. It first introduced the background of open network equipment from the aspects of the open source community and industry status quo, and further analyzes the key technologies involved in the open network equipment research and development, finally it expounds China Unicom's innovative practice of capability opening in the field of 5G bearer network access routers, and the follow-up research and development direction is prospected.

Keywords:

5G bearer; Open network device; Network operating system; SDN

引用格式:徐博华,庞冉,张学茹,等. 电信级开放网络设备关键技术与研发实践[J]. 邮电设计技术, 2022(4): 10-15.

1 概述

专用网络设备的能力开放是软件定义网络(SDN)技术在专用设备层面的自然延伸,有别于网络功能虚拟化(NFV),网络设备能力开放并不追求将网络功能部署在通用服务器上,而是强调通过一系列软硬件解耦的技术手段,在提升软件功能灵活性的同时兼顾专用硬件的高转发性能,因此更加适用于交换机、路由器、传输设备等需要大吞吐量的应用场景。推动专用

网络设备能力开放的过程可分为灰盒和白盒2个阶段^[1]。灰盒^[2]是指网络运营方对设备提出功能和性能要求,并定义统一接口规范,供应商遵循网络运营方提出的规范进行设备开发和接口开放。白盒^[3-4]则是一种更为彻底的解耦模式,网络运营方基于开源或商用的网络设备操作系统(Network Operate System, NOS)构建自己的系统平台及上层应用,并通过标准接口适配白盒设备硬件。白盒模式要求网络运营方具有较强的自主研发能力,同时也依赖产业链各方的深度合作。

数据中心白盒交换机是开放式网络设备成功应用的典型案例,阿里巴巴基于开源Sonic操作系统增量

基金项目:国家重点研发计划(2020YFB1804701)

收稿日期:2022-02-28

开发了 AliNOS, 并实现了数据中心白盒交换机的规模部署。此外, 随着 5G 网络的规模建设, 开放式城域网网络接入路由器也逐渐成为近年来业界关注的热点^[5]。

目前电信运营商在网络领域的创新研发工作主要包含协议的标准化、新技术研究的需求输入等。网络设备研发是新技术研究的直接落地环节, 是网络建设与运营管理的前置环节, 而该环节目前完全依赖于设备供应商, 研发符合电信级高可靠性要求的开放的网络设备具有重要意义。国内 5G 承载网络的建设在城域接入层引入了每年 10 万端级的设备新建需求, 且相较于核心汇聚层设备而言, 其性能指标要求较低, 有利于开展自研工作和生态合作。本文重点围绕 5G 承载网络接入层的场景, 第 2 章对开放网络设备的上中下游产业链和相关开源生态进行了分析, 第 3 章总结了开放网络设备所涉及的关键技术, 第 4 章介绍了中国联通在 5G 承载网络接入层开展的开放网络设备研发实践, 最后对下一步的工作方向进行了展望。

2 开放网络设备产业链分析

2.1 产业链上游

开放网络设备产业链上游包含 CPU、FPGA、CPLD、存储器件、背板连接器、时钟芯片、交换芯片、网络处理器 (Network Processor, NP) 等元器件, 以及电源、风扇等外围器件。其中交换芯片和 NP 是网络设备的核心元器件^[6], 负责将数据流量进行线速转发, 转发过程中需要实现数据接收、转发决策、报文存储、报文编辑等操作, 高速处理包括二层交换、三层路由、ACL、QoS 以及组播等在内的多种业务。交换芯片和 NP 作为网络设备的专用元器件, 是其区别于通用设备的主要特征。

在交换芯片方面, 高端交换芯片市场的主要供应方为美国博通公司。博通的交换芯片产品分为 StrataXGS Trident 系列、StrataXGS Tomahawk 系列和 StrataDNX 系列, 分别适用于企业园区网络、超大型数据中心和电信运营商网络等不同场景。在 StrataDNX 系列下, BCM88480 产品^[7]适用于 5G 承载网络接入层设备, 其最大交换容量为 800 Gbit/s, 同时支持 FlexE、Segment routing、1588v2 等新一代承载网络特性。此外 BCM88280 同样适用于 5G 承载接入层设备, 提供 360 Gbit/s 的交换容量, 用户可根据组网场景进行选择。

国内方面, 盛科网络公司的 CTC7132 产品也可满足 5G 承载接入层对端口、表项、特性的需求, 同时内置

ARM Dual Core A53 CPU, 可以有效降低整机硬件成本, 其新一代产品 CTC8180 新增了对 FlexE 和 SRv6 特性的支持。

2.2 产业链中游

开放网络设备的产业链中游是整机硬件。由于网络设备属于专用设备, 没有 PC 机和服务器等通用设备的开放性好, 因此也更加需要开源生态的引导。OCP 社区为推动数据中心白盒交换机的发展做出了重要贡献, 通过制定 OCP 规范, 开放硬件的设计方案和驱动程序, 得到了众多 ODM 厂商的支持。TIP 是由电信运营商主导的开放网络设备社区, 其主导的分解式移动蜂窝基站网关 (Disaggregated Cell Site Gateway, DCSG) 搭载开源或商用的 NOS 系统, 已在 AT&T、Telefonica、BT 等国外运营商的移动回传网络中得到了一定规模的部署。国内方面, ODCC 先后成立了凤凰项目、开放网络硬件项目和 S3IP 项目^[8], 有效推动了国内交换机白盒化的发展。

在 5G 承载接入层设备领域, 目前国内设备选型以插卡式设备为主, 通常配备双主控卡提高单台设备的稳定性, 并配备不同类型的接口卡来提高组网的灵活性。然而插卡式设备的硬件复杂度相对较高, 白盒化的产业生态还不成熟, 有赖于国内运营商进行积极引导。近期, CCSA 下设的云网智联产业推进委员会 (TC610) 成立了开放网络设备工作组, 将重点推进运营商场景下网络设备的架构开放和硬件标准化。

2.3 产业链下游

开放网络设备的产业链下游是网络所有者, 包含电信运营商、云商、政企用户等, 通过网络设备操作系统的能力开放赋予最终用户定制系统和上层网络应用的权利。开放网络设备操作系统包含商业操作系统和开源操作系统 2 类。商用操作系统主要以 DriveNets、Cumulus network、Big Switch、IP Infusion 等国外厂商的产品为代表。

在全球市场, 白盒交换机操作系统以 Microsoft、Facebook 和 Google 主导的开源项目为主。Sonic 是微软在 2016 年 3 月的 OCP 峰会上正式发布的一款开源网络操作系统^[9]。借助其 2015 年向 OCP 贡献的 SAI 以及基于 docker 的设计方式, Sonic 的生态逐渐明朗, 吸引了众多公司参与, 大型互联网公司以及传统设备厂商如 Alibaba、Tencent、Baidu、Cisco、Juniper、Docker 公司向开源社区贡献代码或提供合作支持, 大型芯片厂商如 Broadcom、Barefoot、Marvell、Centec 等提供了对

ASIC的支持。经过多轮迭代, Sonic已经拥有了丰富的业务功能以及优良的平台特性, 不仅满足数据中心的部署要求, 更是吞噬着部分传统交换机的市场。

Stratum是开源网络交换机操作系统, 目前由开源社区ONF负责运营, 是一个实现了软件定义的数据平面平台, 被定义为符合SDN2.0概念的项目^[10]。Google是Stratum首个版本的贡献者, 也是首个在生产网络中部署的企业。除了谷歌, 包括腾讯、中国联通在内的许多运营商、CT厂商、ODM厂商和芯片厂商都加入了该项目当中。Stratum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。Stratum公开了一组下一代SDN接口, 包括P4 Runtime和OpenConfig, 实现了转发设备的可互换性和转发行为的可编程性。

2017年11月, AT&T发布了致力于开发白盒交换机操作系统的分布式网络操作系统dNOS项目。2018年3月在开放网络峰会上AT&T宣布将dNOS正式托管给Linux基金会, 并且将项目名称改为DANOS^[11]。AT&T已经在自己的IPRAN场景中部署了使用DANOS的接入路由器, 相较于Sonic和Stratum这2种诞生于数据中心的网络操作系统而言, DANOS更适合运营商承载网络的应用场景。2019年底, DANOS进行了操作系统框架的开源, 但功能模块并未全部开源, 且由于其商业模式并不清晰, 产业引导的效果不理想, DANOS项目的前景不甚明朗。表1从产业生态的角度比较了Sonic和DANOS操作系统的现状。

表1 Sonic和DANOS产业生态现状对比

对比项	Sonic	DANOS
特性集	少: 仅关注数据中心网络场景	多: 支持数据中心和运营商网络特性集
开源程度	高: OCP最活跃的项目, 代码全部开源	低: 开源程度低, 核心代码未开源
硬件支持	多: 对BCM、Marvell、Barefoot、Centec等众多交换芯片提供支持	少: 仅支持基于BCM & DPDK/X86的驱动
软件贡献	多: 微软、BAT等均贡献了源代码, BCM等芯片厂商也做了贡献	少: AT&T主导
可扩展性	高: 基于Linux, 网络应用基于docker容器构建, 方便系统应用升级	高: 基于Linux, 在操作系统级别管理数据, 并通知驱动程序进行处理
对框式设备的适配	进行中	尚不明确
商业模式	免费	尚不明确

3 开放网络设备关键技术

3.1 统一北向接口

在黑盒网络设备模式下, 传统设备供应商提供的SDN管控接口存在很多私有数据模型, 导致运营商每引入一家新的设备供应商均需与自研的SDN管控系统进行特异性适配, 开发工作量大, 设备上线周期长。在开放网络设备模式下, 由网络运营商定义设备北向接口规范, 设备操作系统集成标准的北向接口引擎与SDN管控系统对接, 能有效避免接口重复适配问题。

3.2 操作系统架构开放

操作系统架构开放是指系统能够以灵活的方式集成网络应用, 网络应用本身与系统解耦。由于网络设备的硬件资源是有限的, 因此通过部署虚机的方式对网络应用进行隔离并不适用。以Sonic为代表的操作系统使用了轻量级虚拟化技术, 不仅很好地解决了网络应用在操作系统上的隔离和调度问题, 同时能够节省大量专门用于系统开销的硬件资源。

如图1所示, Sonic的网络应用能力基于docker容器构建, 具备较好的扩展性, 无论是新增加功能, 改变一个容器, 还是开发新的特性, 都只需要对局部做出改变, 方便在生产环境实现不停机部署或升级应用, 而不需更换交换机镜像, 最大限度简化运维人员的管理和配置工作量。Sonic将每个模块放置在独立的docker容器中, 在一类功能组件内保持了较高的内聚性, 减少了相互分离组件之间的耦合, 因而这些组件中每一个都可以完全独立编写, 而不需了解与底层平台基础系统的交互细节。

3.3 软硬件解耦

为实现硬件标准化和软硬件解耦, 网络操作系统的基础软件平台层需要具备较好的可移植性, ONL和ONIE是具备较好硬件生态支持的基础系统平台。ONL是一种用于裸机交换机的Linux发行版, 是Debian Linux的分支, 包括一些系统基础的驱动功能, 如: USB、GPIO、Flash、PCIe、Serial、RTC、DMA等, 操作系统厂商可以基于ONL来开发网络功能的特性, 不同白盒设备通过对ONLP API的支持实现软硬件的解耦。ONIE是一个网络操作系统安装程序, 实际是一个基于Linux的小型操作系统, 可以在交换机上启动并发现本地网络上可用的安装程序映像, 同时能够将合适的映像传输到交换机, 然后提供一个安装环境, 以便安装程序可以将NOS加载到交换机。ONL操作系统为上层提供基础性平台, 通过ONIE/bootloader加载NOS软件^[12]。

3.4 数据平面可编程

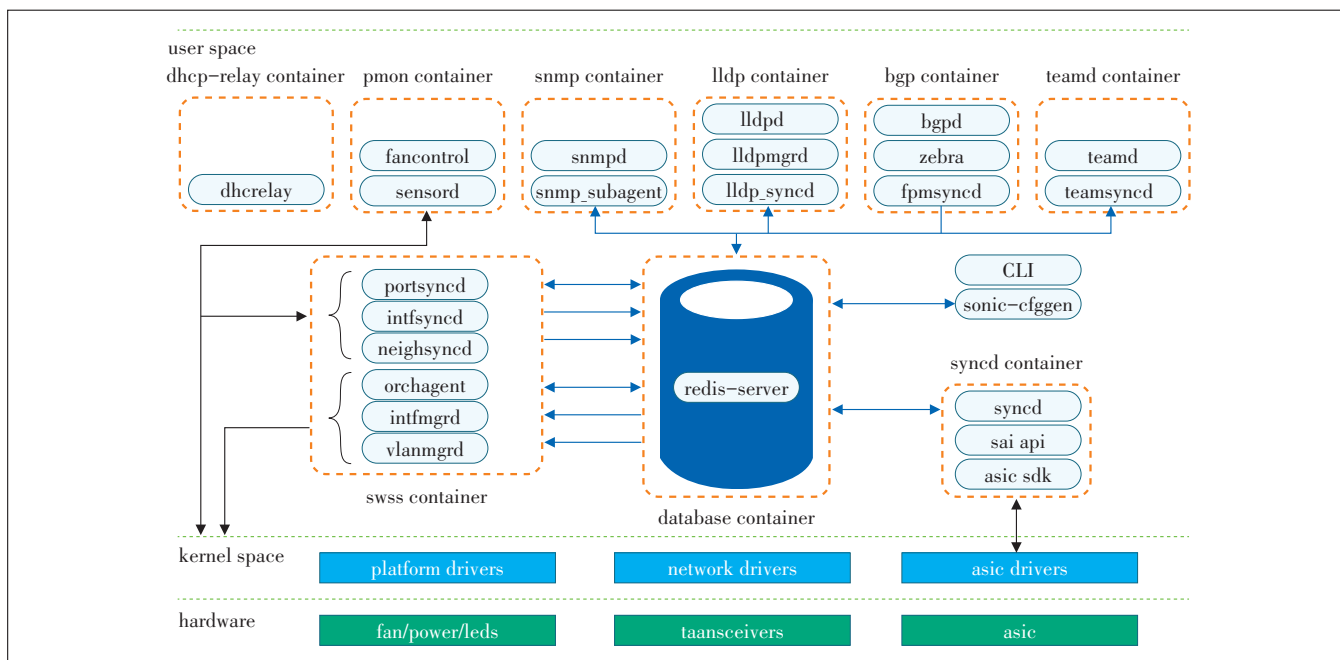


图1 Sonic 开源操作系统架构

由于长久以来网络行业一直被封闭的、专有的硬件和软件所主导,同时标准化的需求导致协议设计的灵活性逐渐降低,网络领域的创新步伐相对缓慢^[13]。数据平面可编程技术通过创新硬件载体和数据平面高级编程语言,使最终用户能够对数据包在交换芯片的转发行为进行自定义。P4是数据平面可编程语言的代表,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^[14],通过 P4 编程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^[15]。

4 开放网络设备研发实践

目前的开源操作系统尚无适用于 5G 承载网络场景的方案,为推进开放网络设备在运营商网络场景下的部署,中国联通研究院在开源社区成熟解决方案的基础上,研发了主要面向电信运营商网络场景的开放网络设备操作系统 CUNOS。CUNOS 的软件系统分层架构如图 2 所示。

CUNOS 系统应用层提供了与外部 OMC 系统进行通信的接口,接受 SDN 管控系统对开放网络设备的统一纳管。应用层子系统内部主要包括 DC 模块、DB 模块和 NESVC 模块。其中 DC 模块完成数据的接收处

理,DB 模块用于配置数据的存储和状态数据的组织,NESVC 模块对于数据进行表校验、表间校验、数据下发以及异常回滚等操作。同时,结合 IETF 和 Openconfig 的相关标准,匹配 5G 承载现网技术方案,制定了中国联通白盒域 Yang 模型标准,并完成了与管控系统的对接测试。

在 CUNOS 系统业务层,需要根据应用层需求完成相应的业务逻辑功能,如模块启动运行、模块参数设置、模块资源分配、协议启动、协议参数设置、协议状态运行、模块和协议数据维护等。业务层除了根据应用层需求完成相应的逻辑功能外,还会为应用层提供多方面的数据信息,如告警、性能统计、各个模块和协议的运行状态等。CUNOS 针对 FRR 开源协议栈进行了二次开发与集成,同时集成了部分商用协议栈模块,使 CUNOS 能够很好地满足 5G 承载网络特性需求。

CUNOS 系统的中间层包含中间适配层和硬件驱动层。中间适配层完成上下层功能之间的转换适配,将不兼容的参数、数据、行为动作根据实际情况进行融合拉通,使用户的需求功能最终可以落到具体的硬件组件上完成执行。中间适配层主要分为板卡适配、端口映射、业务层模块适配、报文收发、中断适配、SAI 接口调用等。其中板卡适配主要完成对于白盒系统支持的硬件板卡进行启动、初始化、运行、卸载、复位、状态获取和板卡属性检测管理等工作。端口映射主

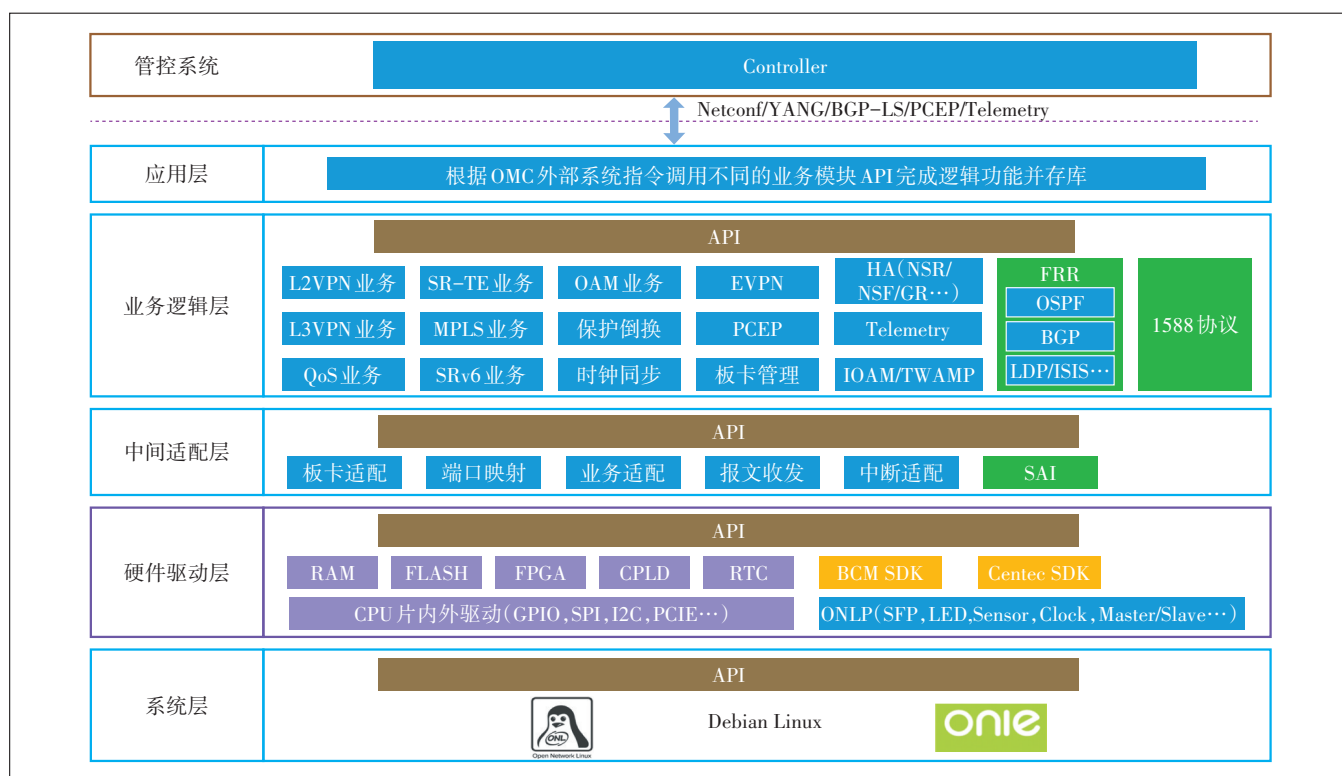


图2 CUNOS操作系统软件分层架构

要完成对上层逻辑端口与底层芯片物理接口之间的映射转换。报文收发和中断主要完成协议报文或者上下CPU的报文的接收与发送。

SAI接口主要对硬件驱动层芯片的SDK进行屏蔽抽象,使得中间适配层针对不同厂家的芯片都可以完成功能特性的下发适配工作。SAI本身是OCP推出的开源项目,但其主要特性集是面向数据中心交换机的场景需求。为满足5G承载网络特性需求,CUNOS针对5G承载需求新增SAI开发模块,如MPLS L2 VPN/L3 VPN、BFD、PTP/1588、TWAMP等,同时对OCP社区已有的SAI模块针对5G承载需求进行了增强,如ACL、HASH、ECMP、QoS、virtual route、routerinterface等(见图3)。

硬件驱动层主要是对硬件组件进行软件驱动,使其可以按照指定需求完成功能动作,包括对RAM、FLASH、FPGA、CPLD、RTC、ASIC芯片、CPU内外部总线的驱动软件功能,以及OCP的ONLP软件抽象功能(对SFP、LED、Sensor、Clock等)。在基础系统层,CUNOS采用了ONL作为Base OS,同时搭配ONIE对上层NOS进行加载,确保CUNOS在开放网络设备上的可移植性。

在硬件设计方面,国内5G承载接入层设备形态通

常为双主控卡+多种类型接口卡的框式设备,而在白盒路由器生态中此类设备形态较少。为此,中国联通研究院联合国内ODM厂商设计了一款基于盛科网络公司CTC7132交换芯片的框式路由器。整机采用框式设计,主控、电源等功能模块提供1+1冗余保护,支持11个业务线卡槽位,整机交换容量400G,IO接口带宽440G,可支持4路25G/10G,4路10G/GE,1路100G/2路50G等多种接口板类型。该白盒路由器支持标准ONLP接口功能,同时针对框式设备进行了ONLP API的扩展定义,主要包括物理层的时钟同步以及主备同步相关功能的抽象。

5 总结与展望

目前,CUNOS操作系统已与白盒框式路由器完成了软硬件的系统集成,通过了实验室测试并进行了试点验证测试。

在算网一体发展趋势下,传统网络设备正快速向ICT融合的方向演进,对开放网络设备的需求愈发强烈。CUNOS操作系统将进一步开放系统上层架构,以轻量级虚拟化的能力集成更多的第三方网络上层应用。同时,在白盒框式路由器上可集成算力增强型的接口卡,面向中小型企业提供就近的边缘算力服务。

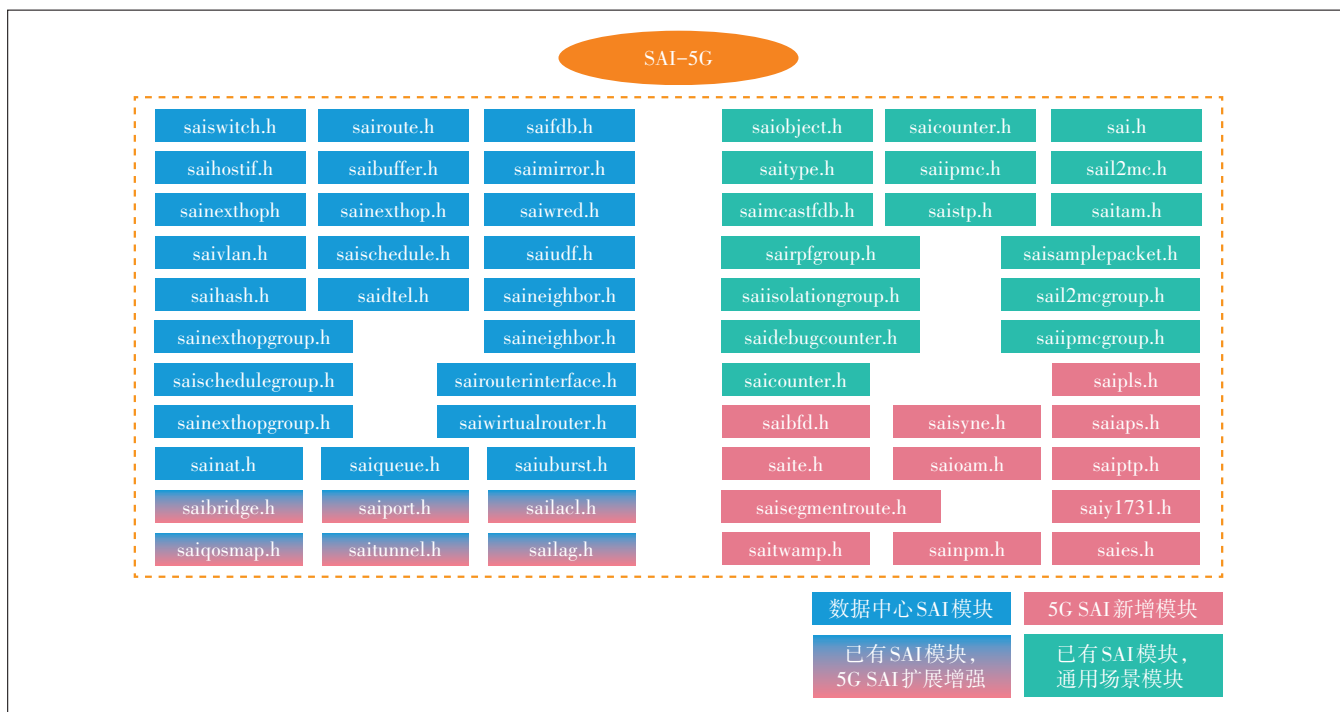


图3 5G承载SAI接口特性集

后续将依托白盒设备的开放特性,与产业链合作伙伴进行联合创新,盘活网络边缘设备能力,为差异化的业务需求提供更加灵活的解决方案。

参考文献:

[1] 梅雅鑫,舒文琼. 网络“黑”“白”之争[J]. 通信世界,2018(29):13-15.
 [2] 胡骞,赵国永,霍晓莉,等. 灰盒光传输设备的SDN管控技术研究及应用[J]. 电信科学,2021,37(12):93-100.
 [3] 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(33):32-33.
 [4] 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(29):20-21.
 [5] Telecom Infra Project. DCSG 概览[EB/OL]. [2022-02-22]. https://www.edge-core.com/cn/solution-inquiry.php?cls=30&id=60.
 [6] 第5届未来网络发展大会组委会. 未来网络白皮书——白盒交换机技术白皮书(2021版)[EB/OL]. [2022-02-22]. https://www.sdnlab.com/resource/25093.html.
 [7] BROADCOM. BCM88480 800- Gbit/s integrated packet processor and traffic manager single-chip device [EB/OL]. [2022-02-22]. https://docs. broadcom. com/doc/BCM88480-800Gbs-PP-and-Traffic-Manager.
 [8] 开放数据中心委员会. S3IP项目(Simplified Switch System Integration Program)[EB/OL]. [2022-02-22]. http://www.opendatacenter.cn/work-group/p-1410875092893126657.html.
 [9] Open Compute Project. Networking/SONiC [EB/OL]. [2022-02-

22]. https://www.opencompute.org/wiki/Networking/SONiC.
 [10] The Open Networking Foundation. Enabling the era of next generation SDN[EB/OL]. [2022-02-22]. https://opennetworking.org/stratum/.
 [11] The Linux Foundation. A unified network operating system driving innovation across the networking stack [EB/OL]. [2022-02-22]. https://www.danosproject.org/.
 [12] Open Compute Project. Open network Linux [EB/OL]. [2022-02-22]. http://www.opennetlinux.org/.
 [13] KFOURY E F, CRICHIGNO J, BOU-HARB E. An exhaustive survey on P4 programmable data plane switches: taxonomy, applications, challenges, and future trends [J]. IEEE access, 2021, 9: 87094-87155.
 [14] HE B H, DONG L G, XU T J, et al. Research on network programming language and policy conflicts for SDN [J]. Concurrency and Computation Practice and Experience, 2017, 29(19): e4218.
 [15] 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.

作者简介:

徐博华,毕业于北京交通大学,工程师,硕士,主要从事IP新技术研究和新型网络设备研发工作;庞冉,毕业于北京邮电大学,高级工程师,硕士,主要从事IP技术研究、IP网络与设备研究及应用相关工作;张学茹,毕业于北京邮电大学,高级工程师,硕士,主要从事IP技术研究、IP网络与设备研究及应用相关工作;王泽林,毕业于西北工业大学,高级工程师,硕士,主要从事IP技术研究、IP网络与设备研究及应用相关工作;曹畅,高级工程师,博士,主要专业领域为IP网络宽带通信、SDN/NFV、新一代网络编排技术等。