

基于自主安全的云数据中心网络技术探索与解决方案研究

张华洪

(迈普通信技术股份有限公司, 四川 成都 610094)

摘要: 基于中国互联网的发展, 以及国家新基建的政策导向, 数据中心正在蓬勃发展, 数据中心网络也在从 10G、40G 网络向 25G、100G 网络过渡, 甚至一些头部企业已经在尝试 100G、400G 网络。同时, 随着中美贸易摩擦日渐升级, 卡脖子的现象日趋严重, 能否基于全自主元器件、网络设备和技术构建云数据中心网络是一个重要课题。云数据中心网络通常包括 SDN 控制器、交换机、安全设备、网络分流设备、虚拟交换机 vSwitch、智能网卡等相关组件, 这些组件的硬件和软件都需要是国产自主的, 才能摆脱对国外技术的依赖。重点论述自主云数据中心网络的技术特点、相关组件自主化布局的可行性以及自主化的方法和方案选择。

关键词: 云数据中心网络; 自主创新; 自主化方案选择

中图分类号: TN918

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.212280

中文引用格式: 张华洪. 基于自主安全的云数据中心网络技术探索与解决方案研究[J]. 电子技术应用, 2021, 47(12): 47-50.

英文引用格式: Zhang Huahong. Research on network technology exploration and solutions for independent cloud data centers of large domestic enterprises[J]. Application of Electronic Technique, 2021, 47(12): 47-50.

Research on network technology exploration and solutions for independent cloud data centers of large domestic enterprises

Zhang Huahong

(Maipu Communication Technology Co., Ltd., Chengdu 610094, China)

Abstract: Based on the development of China's Internet and the policy orientation of the country's new infrastructure, data centers are booming, and data center networks are also transitioning from 10G and 40G networks to 25G and 100G networks. Even some leading companies are already experimenting with 100G and 400G networks. At the same time, with the escalating trade friction between China and the United States, the phenomenon of neck stuck is becoming more and more serious. Whether it is possible to build a cloud data center network based on fully autonomous chips and technology is an important issue. Cloud data center networks usually include SDN controllers, switches, security devices, network offloading devices, virtual switches vSwitch, smart network cards and other related components. The hardware and software of these components need to be domestically independent to get rid of the dependence on foreign technology. This article focuses on the technical characteristics of the autonomous data center, the feasibility of the autonomous layout of related components, and the choice of autonomous methods and solutions.

Key words: cloud data center network; independent innovation; independent plan selection

0 引言

基于中国互联网的发展以及国家新基建的政策导向, 数据中心正在蓬勃发展, 数据中心网络也在从 10G、40G 网络向 25G、100G 网络过渡, 甚至一些头部企业已经在尝试 100G、400G 网络。同时, 随着中美贸易摩擦日渐升级, 卡脖子的现象日趋严重, 能否基于全自主元器件、网络设备和技术构建云数据中心网络是一个重要课题。云数据中心网络通常包括 SDN 控制器、交换机、安全设备、网络分流设备、虚拟交换机 vSwitch、智能网卡等相关组件, 这些组件的硬件和软件都需要是国产自主的, 才能摆脱对国外技术的依赖。本文重点论述自主云

数据中心的技术特点、相关组件自主化布局的可行性以及自主化的方法和方案选择。

1 云数据中心网络架构

自主云数据中心网络架构如图 1 所示。以自主研发的全新架构网络控制器为核心, 向上联动云平台, 提供网络业务的统一编排和呈现; 向下管理网络设备、虚拟交换机(vSwitch)、安全设备、网络分流器、智能网卡, 自动根据云平台的指令打通网络连接, 构建基于 VXLAN 的 overlay 网络基础设施; 同时, 网络控制器整合人工智能技术和网络流量分析技术可以更为智能地管理网络以及呈现网络流量的实时情况。该系统包含以下几个关

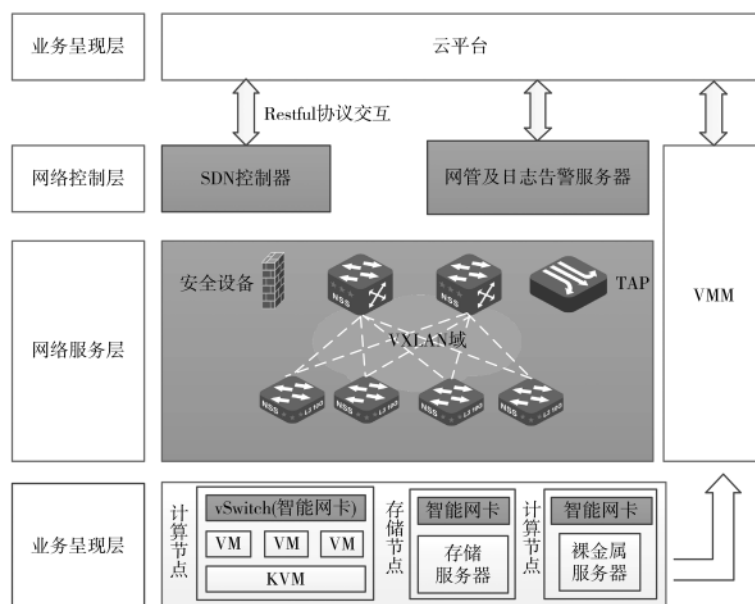


图1 云数据中心网络架构图

键层次。

业务呈现层: 主要指云平台, 包括第三方云平台和 OpenStack 等开源云平台, 负责计算资源管理和网络业务的统一编排。其不属于云数据中心网络, 只和网络有对接和互通关系。

网络控制层: 包括 SDN 网络控制器等控制节点, 主要管理网络设备、vSwitch、安全设备、网络分流器等网络设备, 并和云平台形成联动, 接收云平台的编排指令, 向下配置网络设备和 vSwitch。

网络服务层: 基于交换机、安全设备、网络分流器等网络设备, underlay、overlay 等技术构建的网络服务层, 负责网络业务转发等。

计算接入层: 基于服务器、智能网卡等产品构建的技术接入层, 可以运行虚拟机、容器、裸金属服务等用户业务。其中智能网卡能够协助卸载运行 vSwitch 等软件功能, 以增强处理性能以及降低服务器开销。

2 云数据中心网络关键技术特征

自主云数据中心网络需要具备智能网卡技术(Virtio、OVS 卸载、安全业务卸载、SR-IOV、裸金属服务器支撑)、设备三平面安全可信技术、云计算中心网络技术(网络虚拟化、大二层网络、M-LAG)、SDN 技术、云网一体化技术、智能网络技术、RoCE 及无损以太网技术、网络流量透视和安全监管技术等信创新一代网络关键技术。通过有效地将传统网络设备、网络控制器、智能网卡、网络分流器等资产和组件有机整合, 为客户应用需求提供安全可控的新一代网络技术整体方案, 从而在芯片、软件、技术、网络、应用等方面打造创新生态系统, 形成可持续的竞争力。

2.1 智能网卡技术

伴随着云计算、大数据、AI 及机器学习的发展, 云数

据中心服务器性能压力越来越大, 而基于智能网卡来完成裸金属服务器的部署、网络卸载、特定业务卸载的诉求越来越强烈。智能网卡技术主要涉及 Virtio 模拟技术、OVS-DPDK 技术、VXLAN/Geneve 隧道卸载技术、存储设备模拟技术、GPU 设备模拟技术、国密加解密卸载技术、SR-IOV 技术、CT(Connection Tracking) 硬件卸载技术等。围绕裸金属服务器部署解决方案、解决国产服务器性能不足问题以及各种业务功能卸载等应用, 需要在体系结构上与服务器操作系统、服务器虚拟化、容器等技术结合, 以及和交换机等外部硬件设备联动, 形成端到端的整网网络优化能力, 提高云数据中心整体效能, 如图 2 所示。

从硬件架构上讲, 智能网卡主要基于国产 CPU 和 FPGA 构建。其中 CPU 用于运行智能网卡管理、控制及核心业务软件系统, 称之为 CCP(Card-Carried Processor)。FPGA 系统作为多业务卸载引擎(Multi-Service Offload Engine, MSOE)为智能网卡提供内、外部通信接口以及业务卸载、网络卸载、存储卸载等功能。另外, 由于当前国产 CPU 的发展水平要远高于国产 FPGA 的发展水平, 这里需要有一个精巧的设计, 就是需要在 FPGA 中实现 FerryMan 模块, 用于搬运那些由于 FPGA 资源或性能限制, 暂时无法卸载到 MSOE 中的业务, 暂时用 CPU 实现。等后续 FPGA 的能力上来了, 可以逐步再往 FPGA 中做, 但是对外的业务呈现是完全一致的, 不会让用户在功能上有感知, 只是在升级特定版本后性能等方面有较大提升。

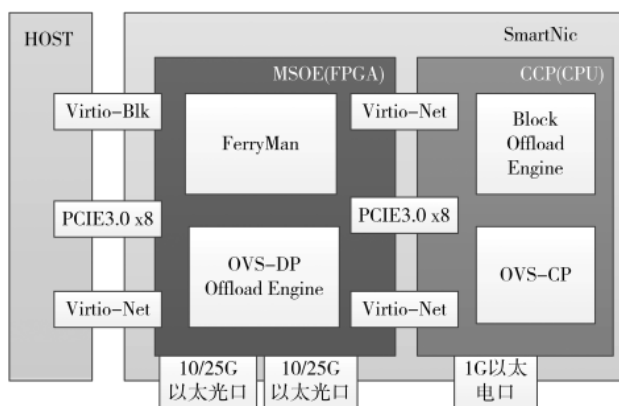


图2 端到端的整网网络优化能力架构

采用国产和自主元器件可以解决供应链安全, 但是不代表应用安全, 所以需要添加国密加解密算法以及对应的安全机制, 使得无论是网络数据还是存储数据都能够进行国密加解密。另外基于国标的 TPCM 可实现智能网卡的可信度量, 还可以度量服务器系统的 BIOS、BMC、操作系统组件。

2.2 SDN 及云网一体化技术研究

随着 SDN 技术的持续发展,SDN 网络控制器已经不仅仅限于流表的学习下发,而是整个网络的调度中心,其向上对接联动云平台,向下管理物理交换机、虚拟交换机、网络防火墙、负载均衡器、网络分流器等各种网络部件,完成云网一体化联动,使得整个云数据中心网络能够根据用户意图自动完成部署和交付。SDN 控制器需要支持云平台联动对接、整网安全策略管控、网络拓扑三网互视、一网多云技术(一套 SDN 网络对接多套云平台)等关键难点技术,达到网络高度自动化的效果。同时全套 SDN 组件需要支持安装在国产服务器上以及支持国产操作系统(比如麒麟操作系统)的适配。

2.3 云数据中心网络技术研究

云数据中心当前主要涉及网络虚拟化、大规模二层网络技术、M-LAG 双规接入技术、服务链技术、微分段技术、虚拟机流量导出技术等。网络虚拟化包括 1:N 虚拟化、N:1 虚拟化、纵向虚拟化三种网络虚拟化功能,大规模二层网络技术主要包括 TRILL 技术、VXLAN 技术、BGP-EVPN 技术以及分流网络技术。这些技术当前业界基本都是成熟的,但是这里需要强调的是自研和自主,不使用国外知识产权的软件,避免国外关闭支撑带来的技术风险。

另外,基于国产交换芯片以及国产智能网卡可以构建属于国产自主的端到端网络应用,比如基于交换机和智能网络共同配合的端到端 INT(Inband Telemetry)可视化功能,端到端 X-PIPE(硬管道确定性网络保障技术),使得用户数据能够得到服务器到服务器的端到端可视以及传输保障,如图 3 所示。

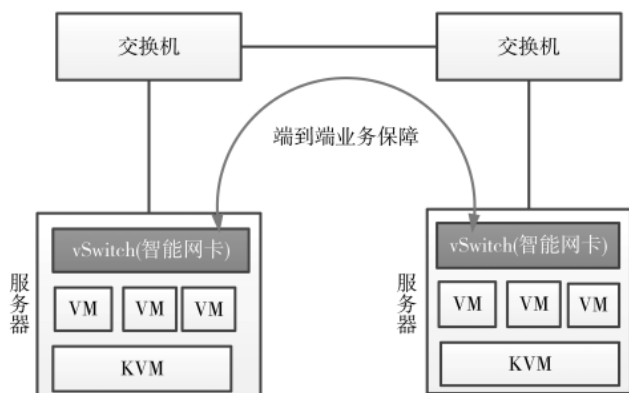


图 3 端到端业务保障

X-PIPE 技术是一种借鉴于时间敏感网络(Time Sensitive Network, TSN)的硬管道技术,其将以太网中的数据区分为高优先级报文和低优先级报文,并且能够保证高优先级报文低延迟、低抖动、零丢包,使得高优先级业务得到绝对的网络传输保障。而端到端 X-PIPE 可以实现从网卡侧就区分优先级并修改帧前导以标记报文优先级,后续所有经过的交换机都基于此优先级进行转发

保障,以此达到端到端业务保障效果。这里区别于传统的 QoS 保障技术,X-PIPE 从最原始的帧前导入手,在报文最底层进行优先级标记,不像 QoS 技术只依赖于交换芯片 buffer 在存储转发的时候做优先级调度,如果接收侧或者发送侧本身就已经丢包 QoS 技术无法保障。

2.4 智能网络技术研究

随着网络技术的发展,云数据中心中 underlay、overlay 网络越来越普遍,虚拟机、容器、裸金属服务器、分布式存储等对于网络设备提出了新的要求。需要网络更具备智能,能够智能地联动云平台管理网络业务、智能地调节网络 PFC、ECN 水线,使得网络零丢包低延迟、智能地进行网络维护管理,使得网络故障诊断和恢复无须人工干预。网络高效运维管理是新一代云数据中心网络面临的技术难题,引入机器学习和最新的 AI 算法,赋能数据中心网络,可以使得新一代云数据中心网络维护管理变得自动而高效。

2.5 网络流量透视和安全监管技术研究

新一代云数据中心中,对于网络流量的透视作为网络运维管理、智能网络管理的基础数据来源以及安全监管的要求,成为云数据中心的标准配置。利用端口镜像、流镜像、Sflow、IPFIX、网络分流(TAP)、网络性能分析、虚拟机流量导出、DPI 等技术,将云数据中心中整网流量导出并进行分析透视,形成图形化呈现,可以让用户能够直观感知网络流量情况,让网络从用户很难理解和感知的黑盒变成可以直观呈现的白盒状态。需要针对云数据中心网络和流量特征,针对性优化了 overlay 流量导出分析、虚拟机流量导出分析、TAP 数据预处理、网络性能分析和呈现等方面进行重点优化和突破创新,满足了用户流量透视和安全监管的应用需求。

2.6 RoCE 及无损以太网技术研究

随着云计算的蓬勃发展,物美价廉的分布式存储由于其只需要通用服务器而非专业存储设备,受到了云数据中心的青睐。但是由于 SSD 等存储技术的高速发展,分布式存储环境中,网络成了影响存储性能的最大瓶颈,需要网络具备零丢包、低延迟的能力。另外,高性能计算、分布式 AI 训练等新兴业务应用,也对网络丢包和延迟提出了高要求。要提升网络性能,需要通过多种技术组合实现,包括 RDMA 及 RoCE 技术、PFC 技术、ECN 技术、DC-QCN 技术、ETS 技术、X-PINE 确定性延迟技术,可以使得云数据中心网络达到零丢包以及不受报文大小影响的 $10\ \mu\text{s}$ 以内的确定性延迟效果。需要通过对以太网存储交换机的优化设计,并联动智能网卡,形成了整网零丢包、低延迟、高稳定性的能力,能够支撑云数据中心分布式存储、高性能计算、分布式 AI 训练的应用。并且网络丢包、延迟、稳定性等效果达到业界领先水平。

2.7 安全可信技术研究

网络安全主要包括以下 4 个方面:信息设备安全、

数据安全、内容安全和行为安全。信息系统硬件结构的安全和操作系统的安是信息系统安全的基础,密码、网络安全等技术是关键技术。只有从信息系统的硬件和软件的底层采取安全措施,从整体上采取措施,才能比较有效地确保信息系统的安全。所有设备都需要基于国产芯片和国产操作系统,并且实现可信计算技术与网络通信设备的有机融合,解决网络设备身份识别、网络系统可控接入,网络设备软件系统身份合法性、数据完整性认证,网络设备加载流程可信控制等,从设备、网络系统不同层次解决安全可信需求。

3 关键组件国产自主化可行性论证

国产网络设备主要依赖于 CPU、交换芯片、安全加密芯片、FPGA、可信计算芯片、操作系统等,当前国内已经形成了相应的产业链,如表 1 所示。

表 1 国产关键组件

芯片或系统	国产芯片或系统
CPU	飞腾、龙芯、海光
交换芯片	盛科
安全加密芯片	沐创
FPGA	安路、复旦微
可信计算芯片	可信华泰
操作系统	麒麟操作系统、统信操作系统(UOS)

只要是自研即可,同时需要注意和国产操作系统(如麒麟操作系统、UOS 操作系统)进行适配对接。

防火墙等安全产品,由于其主要依赖的是 CPU,飞腾 CPU 的性能已经能够满足大多数情况应用,如果需要非常高性能处理能力,可以考虑海光 CPU,其 X86 架构和 Intel 公司 CPU 差距不大。同时如果是一些强安全行业,还依赖于安全加密芯片和可信计算芯片,当前沐创、可信华泰等国内厂商已经有成熟产品。

交换机、分流器等产品对于 CPU 依赖不强,飞腾 CPU 已经足够满足要求。但是其对于交换芯片的依赖较高,盛科交换芯片当前能够满足 25G、100G 网络,适应大多数云数据中心建设要求。

操作系统层面,当前麒麟操作系统、UOS 等都是基于 Linux 进行安全强化而来的,也能够满足应用要求以及安全性述求。

上层应用软件,需要各厂商进行自主研发,这部分也是核心竞争力。

4 结论

基于国内 CPU、交换芯片、安全加密芯片等硬件技术以及操作系统、上层应用软件等软件技术的发展,基于国产自主创新的云数据中心生态已经基本成型,虽然在最高端的互联网数据中心还无法对国外产品进行替代,但是大多数云数据中心已经能够完全替代国外芯片及技术产品。

参考文献

[1] 沈昌祥.可信计算平台与安全操作系统[J].网络安全技术与应用,2005(4):8-9.

[2] 沈昌祥.可信计算构筑主动防御的安全体系[J].信息安全与通信保密,2016(6):34.

[3] 曹焱.关于云数据中心网络架构分析及关键技术探讨[J].长江信息通信,2021,34(7):121-123.

[4] 魏喜莲.云数据中心网络安全设备部署研究[J].铁道通信信号,2021,57(4):41-47.

[5] 周洪利.SDN 在云数据中心网络的应用与实现[J].信息技术与信息化,2021(2):173-175.

[6] 刘培振,彭颖颖,吕宗凯.云数据中心网络架构及关键技术[J].中国新通信,2020,22(1).

[7] 杨瀚琛.探讨云数据中心及云平台技术应用[J].电脑知识与技术,2020,16(6):79-81.

[8] 耿延军,王俊,周红亮.云数据中心网络纵深防御研究[J].信息安全与通信保密,2019(7):22-29.

[9] 钟掖,龙玉江,赵威扬,等.基于软件定义网络的电力云数据中心流量控制技术[J].科学技术创新,2021(15):80-81.

[10] 罗鉴,宋晓丽.无损以太网关键技术研究[J].中兴通讯技术,2012,18(4):27-31.

[11] 刘畅,陶云祥,戎胤.智能网卡应用于云网络加速方案研究[J].电信工程技术与标准化,2020,33(8):76-81.

[12] 马潇潇,杨帆,王展,等.智能网卡综述[J].计算机研究与发展,2021.

[13] 马威,姚静波,常永胜,等.国产 CPU 发展的现状与展望[J].集成电路应用,2019,36(4):5-8.

(收稿日期:2021-10-31)

作者简介:

张华洪(1981-),男,本科,工程师,主要研究方向:数据中心网络、智能网络、云计算、嵌入式系统设计、开源网络、物联网。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所