



中国移动 DPU 技术白皮书

(2022 年)

中国移动通信有限公司研究院

2022年7月

主编单位

中国移动通信有限公司研究院

联合编写单位

中国信息通信研究院、中国移动云能力中心、
中兴通讯股份有限公司、华为技术有限公司、英特尔（中国）有限
公司、联想（北京）有限公司、浪潮电子信息产业股份有限公司、
芯启源电子科技有限公司、北京大禹智芯科技有限公司、益思芯科
技（上海）有限公司、新华三技术有限公司

前 言

算力正在逐步成为新时代的核心生产力，算力的发展同步带动了网络的变革，网络作为连接用户、数据与算力的桥梁，是优化算力服务的根基，构建高效、灵活、敏捷的算力基础设施底座，成为达到中国移动算力网络新型基础设施愿景的关键。

作为算力网络创新技术之一，算力卸载统筹虚拟化、数据安全、运维管理等领域，是构建高性能、高可靠云化平台的关键技术。DPU 一方面是实现算力卸载的重要载体，另一方面也是算网一体的初级形态，定位于数据中心继 CPU、GPU 之后的“第三颗主力芯片”，业界需要重点打造和推动 DPU 技术的发展与成熟。

本白皮书希望通过阐明 DPU 发展情况、介绍 DPU 技术架构、分析 DPU 落地所面临的挑战且抛出 DPU 标准化建议，传递中国移动推动 DPU 技术发展的理念，拉通 DPU 完整生态链条，推动 DPU 技术发展成熟。

目 录

目 录.....	1
1. DPU 概述.....	2
1.1 DPU 的背景和概念.....	2
1.2 DPU 产业现状.....	4
2. DPU 技术架构.....	9
2.1 管理系统卸载.....	9
2.2 网络系统卸载.....	10
2.2.1 虚拟化网络.....	10
2.2.2 RDMA.....	13
2.3 存储系统卸载.....	14
2.3.1 存储接口卸载.....	14
2.3.2 存储网络协议栈卸载.....	15
2.4 计算系统卸载.....	16
2.5 安全系统卸载.....	17
2.5.1 虚拟化安全隔离.....	18
2.5.2 网络数据加解密.....	18
3. DPU 技术挑战和推进建议.....	19
3.1 软件系统标准化.....	21
3.2 硬件系统标准化.....	24
4. 展望与产业诉求.....	26
附录：术语、定义和缩略语.....	27

1. DPU 概述

1.1 DPU 的背景和概念

随着工业互联网、人工智能、大数据、5G 等产业快速发展，计算需求呈爆炸式增长，数据中心中的海量数据流动驱动网卡的端口速率从 10G 快速向 25G 甚至 100G 及以上演进，云化基础设施平台基于 CPU 完成网络数据转发的传统模式出现瓶颈。一方面，CPU 串行计算模式难以在高并发网络数据转发工作中发挥最大计算能力；另一方面，网络带宽的增长速度远超 CPU 的算力增长速度，网络数据转发占用的 CPU 资源不断增大。

业界提出的 DPDK（Data Plane Development Kit）加速方案，虽然通过用户态绑核轮询的模式、绕过内核协议栈提升了网络 I/O 数据处理性能，但是在 25G 带宽速率下，大部分数据中心所需 CPU 开销就达到了 25% 甚至更多，随着带宽增长，这种方案的 CPU 开销更加难以忽视。

在这种背景下，智能网卡（SmartNIC）技术走进人们视野并逐渐得到普及。传统网卡在完成数据收发工作外，不处理任何与云化、业务相关的计算任务。而网卡作为串行在网络与计算之间的桥梁，是可以解决计算瓶颈的关键硬件。不同于传统网卡，智能网卡卸载云化网络转发功能，提供高性能网络转发能力，释放 CPU 通用计算资源。同时，提供网络可编程能力，实现网络转发逻辑的定制化，保证了功能的灵活性。智能网卡提供了一种以更高性价比实现网络加速的技术方案，同时也提供了一种新的设计思想。

延续智能网卡卸载计算的设计思想，业界发现可以进一步解决云计算领域的更多问题，如：

- 资源争抢限制：同一计算节点上的云主机实例与虚拟化软件共享计算资源，随着资源需求的提高，资源争抢容易造成服务质量不稳定，尤其在大负载、大流量时 I/O 性能容易出现严重抖动，无法保障稳定的 SLA 体验；

- 计算特性损失：云计算技术的核心是 Intel VT 等硬件辅助虚拟化技术和 KVM 等主流虚拟化系统软件，通过这些技术为用户提供类似物理服务器资源的虚拟机。但是，一方面，虚拟机相比物理机存在一定的性能损失；另一方面，由于客户无法在虚拟机中再次部署虚拟化系统，导致私有云无法更好地利用公有云弹性云主机资源，限制了云主机的使用场景；
- 裸金属管理问题：裸金属可以解决虚拟化带来的计算特性损失问题，但是裸金属 CPU 资源需要全部提供给用户，因此无法在主机 CPU 上运行云平台管理组件，需要通过带外管理流程完成裸金属实例的部署、交付和运维，整个流程与虚拟化场景差异较大，不符合云计算弹性、灵活的原则。此外，裸金属实例在对接远端分布式存储时由于存储客户端需要运行在属于用户的 CPU 上，存储网络暴露存在安全风险。

此外，随着人工智能业务的普及，越来越多的云上 AI 计算任务对网络和存储 I/O 的时延性能提出了更极致的需求，RDMA（Remote Direct Memory Access）和 NVMe（NVM Express）等高性能网络和存储协议在传统网卡架构下难以满足云计算多租户的灵活需求场景。这种背景下，为解决后摩尔时代 I/O 性能瓶颈和虚拟化技术发展限制等诸多问题，DPU 应运而生。

业界对 DPU 概念存在多种解读，如数据处理器（Data Processing Unit）、数据中心处理器（Datacenter Processing Unit）以及以数据为中心处理器（Data-centric Processing Unit）。无论以何种维度进行解读，DPU 从本质上来看，均是一种围绕数据处理提供网络、存储、安全、管理等数据中心基础设施虚拟化服务的专用处理器，基于 ARM/X86 等架构的 CPU 与 ASIC（Application Specific Integrated Circuit）/NP（Network Processor）/FPGA（Field Programmable Gate Array）等专用硬件加速引擎组成的计算架构，形成提供虚拟化功能的实体。作为主机的数据出入口，DPU 在具备标准网卡能力的同时，利用专用硬件完成网络和存储 I/O，释放主机 CPU 算力资源的同时可显著提升 I/O 性能。业务（主机 CPU）与虚拟化软件（DPU）的硬件载体分离，业务与云平台的隔离性以及主机的安全性进一步提高。运行在 DPU 中的管理软件可提供裸金属的云化管理能力，提升裸机业

务灵活性，降低运维难度。

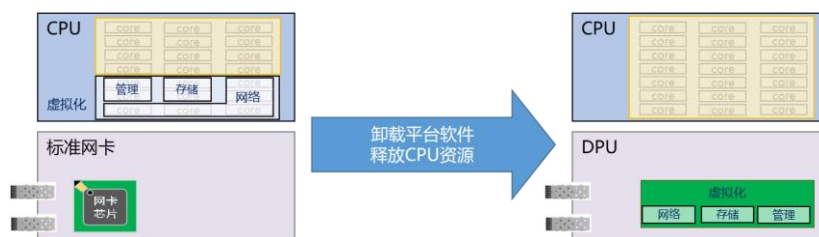


图 1-1 DPU 在云计算场景实现全卸载

相比于传统网卡，DPU 在单部件成本上有所增加，但是 DPU 的引入解放了更高成本的主机 CPU 算力，释放了更多可售卖资源，因此 DPU 引发的架构变革一定程度上提升了整个数据中心资源池的能效成本比和公有云厂商的收益成本比。未来 DPU 可以覆盖越来越多的需求和场景，在这一机遇背景下，产业的爆发成为必然。

1.2 DPU 产业现状

DPU 最早由 Fungible 提出概念，随着 Mellanox 被 NVIDIA 收购，由 BlueField 产品衍生的 DPU 产品概念的影响进一步增强，引发国内外众多厂商密切关注。总体来看，DPU 按产业模式可主要分为商用 DPU 和定制 DPU 两大类：

一、商用 DPU：NVIDIA、Intel 等厂商推出的基于自研芯片的商用 DPU 产品，提供卸载功能和 AI 算力等，云平台开发者配合配套软件栈调用接口即可实现功能模块的卸载和性能加速。

1. NVIDIA DPU

NVIDIA 于 2022 年 4 月发布了第三代 BlueField，并具备持续演进能力，演进路线以集成 AI 算力、增强 ARM 能力以及升级网络速度带宽为主。BlueField 架构本质上将网卡子系统与可编程数据路径、用于加密、压缩等的硬件加速器子系统以及用于控制的 ARM 处理器子系统融合在一起。BlueField-3 DPU 继承了 BlueField-2 DPU 的先进特性，并对其进行了性能加强与功能扩展，在网络业务中专门对 RDMA、连接跟踪、ASAP²（Accelerated Switch and Packet Processing）等网络通信技术做了进一步加强，特别是对时间精度，可在数据中心和边缘之间

进行及时的时钟同步；在存储业务中支持块存储、文件存储、对象存储或者 NVMe 存储的仿真，同时支持数据落盘时的加解密操作。此外，NVIDIA 推出了 DOCA 软件开发包，为开发者提供一个完整、开放的软件平台，支持开发者在 BlueField DPU 上开发网络、存储、安全和管理等虚拟化应用。

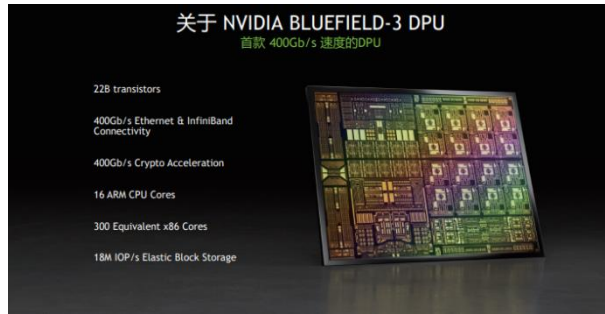


图 1-2 BlueField DPU

2. Intel IPU

Intel IPU（基础设施处理单元，Infrastructure Processor Unit）提供了专用的可编程内核来加速和管理基础设施功能。IPU 与 DPU 的设计理念有很多相似之处，通过将基础设施相关的任务卸载到 IPU，可减少 CPU 开销。

在 2021 年英特尔架构日上，英特尔展示了其首款专用 ASIC IPU Mount Evans 和基于 FPGA 的 IPU 参考平台 Oak Springs Canyon。Mount Evans 融合了英特尔多代 FPGA SmartNIC 的研发经验，提供高性能网络和存储虚拟化卸载以及可编程数据包处理引擎，支持防火墙和虚拟路由等功能。同时延续了英特尔高性能 Quick Assist 技术以及高级加密、压缩加速等技术。

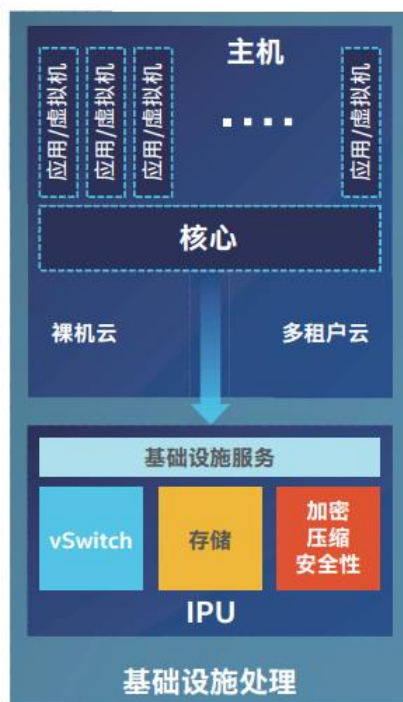


图 1-3 Intel IPU

Intel IPU 配套的基础架构开发套件 IPDK 提供了统一开放 API 来支持业务应用开发和基础设施管理。值得一提的是，Intel 提出，IPDK 是一个开源的、与厂商无关的驱动程序和 API 框架，可管理由 CPU、IPU、DPU 或交换机构建的基础设施。

随着 DPU/IPU 持续爆火，越来越多的厂商相继入局，6 月 21 日，Linux 基金会宣布了新的开放可编程基础设施（OPI）项目，IPDK 与 DOCA 均贡献给 OPI 项目，OPI 能否统一市场上 DPU/IPU 的软件生态仍有待时间检验。

二、定制 DPU：AWS、阿里云等头部云计算厂商推出自研定制 DPU 产品，例如 Nitro 系统和 CIPU（云数据中心专用处理器，Cloud infrastructure Processing Units），结合自研软件系统提供软硬一体化的极致高性能、低成本的云计算服务。

1. AWS Nitro:

Nitro 系统通过软硬件创新，给 AWS 云服务用户带来了更丰富的计算实例和更高性价比选择，已成为 AWS 云服务实例的技术基石。Nitro 系统主要包括三大部分：

- Nitro Hypervisor: 专有硬件上承载 Hypervisor，Hypervisor 不再占用

主机资源不影响主机性能，让主机实现近似裸机服务器的性能表现；

- Nitro Cards: 专有硬件承载存储、网络功能，以及 Nitro 系统控制；
- Nitro 安全芯片: 实现硬件层的安全验证能力。

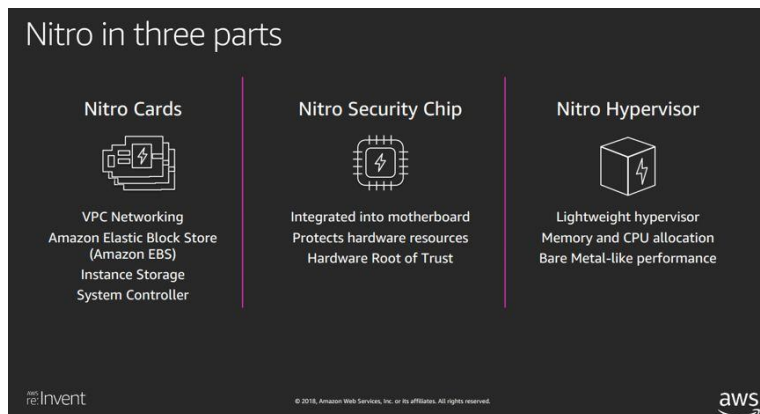


图 1-4 AWS Nitro

2. 阿里云 CIPU:

2017 年，阿里云推出了自研第一代神龙服务器，搭载自研 MoC (Microserver on Chip) 卡，MoC 卡能力伴随神龙架构的演进也在迭代更新：

- 神龙 1.0: 解决上云后如何支持裸机服务的问题，将云化组件尤其是裸金属的管理模块卸载到 MoC 卡，提供裸金属的弹性交付和运维；
- 神龙 2.0: 神龙芯片能力进一步增强，通过构建轻量级 Hypervisor 实现计算虚拟化的卸载，支持虚拟机服务；
- 神龙 3.0: 存储、网络等数据面路径全面通过 DPU 芯片硬件优化，性能大幅提升，可以提供接近裸机的低延时网络；
- 神龙 4.0: 融入弹性 RDMA 能力，让 RDMA 从 HPC 类应用，走向支持通用类计算场景。

以 MoC 卡作为雏形，2022 年 6 月阿里云发布云数据中心专用处理器 CIPU，宣称将成为云时代 IDC 的处理核心。CIPU 向下接入物理的计算、存储、网络资源，快速云化并进行硬件加速；向上接入飞天云操作系统，管控阿里云全球上百万台服务器。以“飞天+CIPU”支撑阿里云实现了软硬件融合的云计算技术体系。

3. 中国移动磐石计算架构和 COCA 算力框架:

中国移动为解决“大云”算力体系的发展瓶颈，重点围绕降低成本、提升性能、增加安全性、提升稳定性等目标，于 2018 年启动自研 DPU 的软硬件生态构建，具体以硬件和软件同步演进思考：

- 磐石算力架构：由自研 HyperCard DPU 和支撑其生命周期的磐石服务器组成，围绕裸金属、云主机、容器等算力需求作统一算力底座支撑；
- COCA（Compute On Chip Architecture）算力框架：面向算力网络技术设施生态的算力卸载、解耦、无损构建目标，力图构建算力的标准化接入生态同时弥补算力与应用间鸿沟。

值得一提的是，中国移动提出 COCA 算力框架是为实现一个软件生态，做到包括自有 HyperCard 在内的业界 DPU 硬件与应用更好的解耦，实现未来云操作系统以 COCA 为介质的 DPU 在数据中心即插即用，这与阿里云 CIPU 的演进思路是相通的。

随着 DPU 产业的迅猛发展，国内涌现了一大批 DPU 初创公司，通过 FPGA、NP、ASIC 等多种硬件路线实现 DPU 芯片的设计，与国内云厂商和各行业数据中心采取定制或联合研发等多种模式，推动 DPU 在数据中心内更广泛地落地，产业竞争日益激烈。

2. DPU 技术架构

从云计算产业来看，主机 Overlay+DPU/智能网卡的网络架构是业界发展的趋势，从中国移动内部来看，算力网络面向智算等业务的建设对云化基础设施性能提出了更高要求，移动云、IT 云、网络云三朵云向虚拟机、容器、裸金属多种服务形态发展，DPU 需求涌现。基于 DPU 的云化架构按需对网络、存储系统进行卸载加速，可以更高性价比获得极致性能。同时，DPU 协同处理云化管理、计算和安全等任务是实现虚拟机、容器、裸金属服务管理方案统一以及云基础设施统一的关键。DPU 云化技术架构是一种云化软件与 DPU 硬件相结合的新型云化架构，以下章节将介绍 DPU 云化技术架构。

2.1 管理系统卸载

OpenStack 是业界主流云平台管理系统，可以为云平台提供裸金属、虚拟机和容器的虚拟资源管理和调度功能。在传统的以 CPU 为中心的云底座架构中，云底座的管理系统与虚拟机或容器共用主机 CPU，由于裸金属实例中 CPU 全部为用户使用资源，因此无法实现裸金属类似虚拟机的弹性管理和交付，限制了不同场景下云底座平台提供计算和服务的能力。

在 DPU 云化架构中，云平台管理组件从主机 CPU 上卸载到 DPU 上运行，将裸金属当作一个大规格的虚拟机来管理和发放，提供与虚拟机服务一致的弹性、灵活、安全的裸金属服务。云平台管理组件对 CPU 的计算性能要求不高，DPU 上 CPU 核的计算性能足以满足需求。

管理系统主要包含虚拟化管理系统和裸金属管理系统两种卸载场景：

- **虚拟化管理系统：**主要为 OpenStack 原生的 Nova-compute、Neutron-agent 等管理组件，这些组件部署到 DPU 上，通过 PCIe 通道与 Host OS 上的虚拟化计算系统交互，完成虚拟机的生命周期管理。相比于传统模式只是管理组件的部署位置以及和 Host OS 交互接口发生了变化，整体管理系统架构保持不变；

- 裸金属管理系统：业界存在 Ironic 和 Nova 两种不同的方式。OpenStack 原生通过 Ironic 组件完成裸金属的生命周期管理，将 Ironic Agent 组件安装到 DPU 上，并对 Cinder 组件做增强开发，配合 Ironic 流程为裸金属挂载云盘；Nova 方式采用 Nova 组件流程，将裸金属当作和虚拟机一样的云主机实例管理，通过 DPU 上部署的 Nova-compute 完成裸金属生命周期管理，与虚拟机管理差异在于无需 libvirt 等计算虚拟化组件。

此外，除上述云主机管理组件外，VNC（Virtual Network Console）、监控脚本、系统日志等相关运维组件同样卸载到 DPU 上，可以降低虚拟化场景下宿主机 CPU 资源开销，也可以为裸金属提供和虚拟机一致的交付和运维体验。

综上，将云平台管理系统卸载到 DPU 上，提升了裸金属管理效率、降低了主机 CPU 管理资源开销、增强了虚拟机性能和稳定性，实现承载虚拟机、容器、裸金属实例的统一云基础设施底座以及提供一致的管理体验。

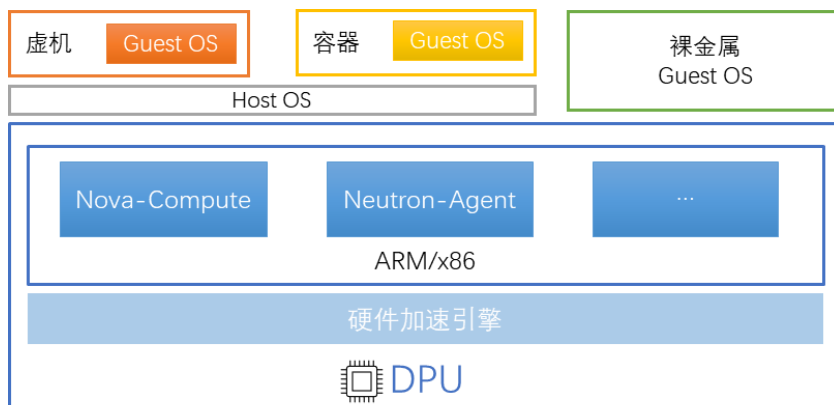


图 2-1 DPU 构建虚拟机、容器、裸金属的统一云基础设施底座

2.2 网络系统卸载

2.2.1 虚拟化网络

虚拟化网络演进是云计算发展的重要推动力。早期业务对虚拟化网络的处理性能要求不高，虚拟网络转发通过软件实现即可满足需求。但是，随着业务的发展与丰富，对虚拟化网络的带宽和时延要求不断提高，通过硬件卸载虚拟化网络成为主流。虚拟化网络硬件卸载主要包含两大部分：网络接口的硬件卸载和虚拟

交换网络的硬件卸载。

1) 网络接口的硬件卸载：

随着虚拟化网络技术的发展，虚拟网络接口多采用 virtio 和 SR-IOV（Single Root I/O Virtualization）两种 I/O 虚拟化技术。SR-IOV 是一种基于硬件的 I/O 虚拟化解决方案，通过分割物理资源实现物理网卡的共享和资源独占，将单个 PCIe 设备虚拟出多个虚拟 PCIe 设备直通到虚拟机中，绕过内核协议栈，获得了极高性能，但存在不支持热迁移以及虚拟机需与网卡驱动定向适配的问题。Virtio 是一种 I/O 半虚拟化技术，采用前后端分离的架构，虚拟机 Guest OS 中安装通用 virtio 驱动，与宿主机中的 virtio 后端进行数据交互。随着性能需求的不断提升，virtio 后端经历了 virtio-net、vhost-net 到 vhost-user 的演进，但是始终存在软件处理网络 I/O 的瓶颈。virtio 技术由于存在控制面与 QEMU 交互，可以实现虚拟机的热迁移，相比于 SR-IOV 更为灵活。从保障计算资源灵活和降低业务上云复杂度的角度来看，virtio 技术更符合云计算的发展需求。

随着虚拟交换网络下沉到 DPU 硬件处理，virtio 后端的实现也要从用户态 vhost-user 模式下沉到 DPU 硬件实现。vDPA 是硬件卸载 virtio 后端的一种开源实现框架，其数据面与 SR-IOV 一样直通到虚拟机中，控制面通过 vhost-user 协议与 QEMU 交互，实现虚拟机热迁移功能。该方案兼顾了高性能与灵活性，成为了网络接口硬件卸载的一种标准方案。从实现方式上看，vDPA 有基于 DPDK 的用户态框架和基于 Linux 内核的内核态框架。由于中国移动选择 OvS-DPDK 的技术路线，DPU 中 vDPA 选择用户态框架实现，架构如图 2-2 所示：

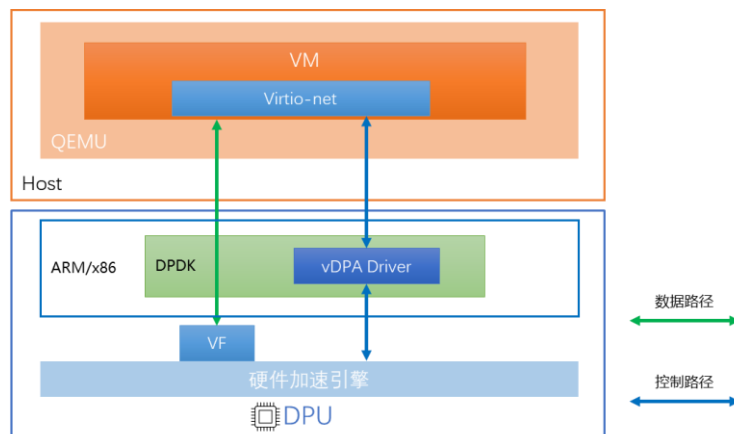


图 2-2 用户态 vDPA 框架

2) 虚拟交换网络的硬件卸载:

在云计算环境中，虚拟交换机 vSwitch 是 Overlay 网络重要组成部分，在当前各种 vSwitch 的实现中，OvS（开放虚拟交换机，Open vSwitch）是业界采用最为广泛的一种虚拟交换的软件。OvS 架构分为控制面和数据面，原生的 OvS 数据面数据处理性能较差，因此出现了 DPDK 数据转发面用户态加速的技术方案，OvS-DPDK 依赖绑 CPU 核轮询的方式完成基于流表 match-action 的转发处理动作。但是随着带宽的提升，需要预留更多的 CPU 资源才能保证转发性能，例如，在中国移动网络云 25G 带宽场景下需要 18%左右的 CPU 资源，在 100G 场景下资源将消耗过大，留给业务的通用算力不断减少。DPU 作为数据入口，是卸载 vSwitch 的最佳选项。

虚拟交换机的控制面软件部署在 DPU 的 CPU 核上，转发面卸载到 DPU 上的硬件加速引擎处理，控制面和转发面之间通过 DPDK 标准 rte_flow 方式进行交互，实现虚拟交换机全卸载，释放主机侧网络转发 CPU 资源开销。如图 2-3，卸载虚拟交换机是 virtio 后端卸载的必要条件，DPU 收到的数据报文在卡上完成 vSwitch 转发处理后，通过 vDPA 框架实现的 virtio-net 后端传递给虚拟机或裸金属内部的标准 virtio-net 前端，数据的传递无需 CPU 的干预，实现更高的网络转发性能。

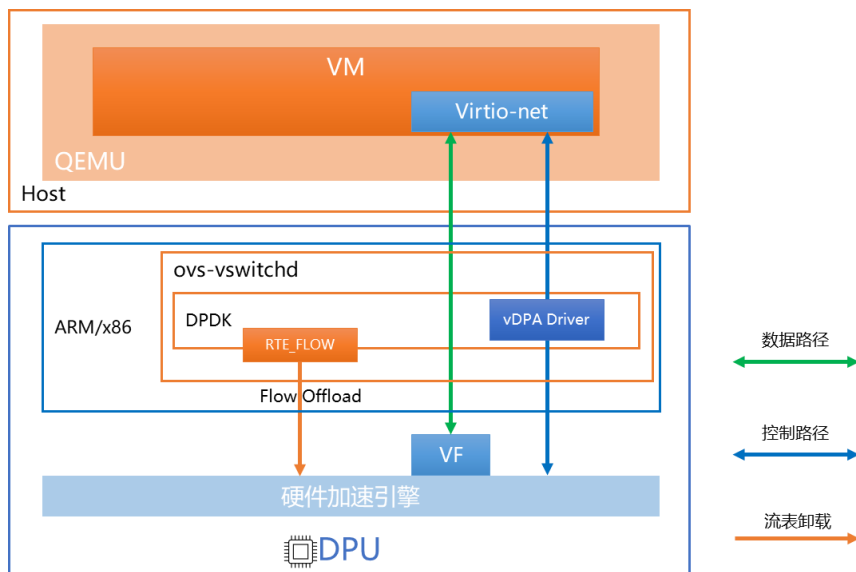


图 2-3 DPU 用户态虚拟交换网络卸载

2.2.2 RDMA

高速存储介质如 NVMe 提供微秒级访问时延，HPC/AI 典型应用依赖低时延网络完成状态同步等操作，对网络时延提出了更高要求，传统毫秒级的 TCP 网络已经无法满足需求。RDMA 技术应运而生，使用内核旁路和零拷贝技术，使能网卡和应用之间的直接数据读写，消除了传统 TCP/IP 协议栈的上下文切换和数据拷贝开销，因此可提供端到端微秒级、以及低 CPU 开销的低时延网络通信。因此在 HPC/AI、Data Base、Web server 等类型的计算节点、高性能存储节点，以及网关节点等得到越来越广泛的应用。

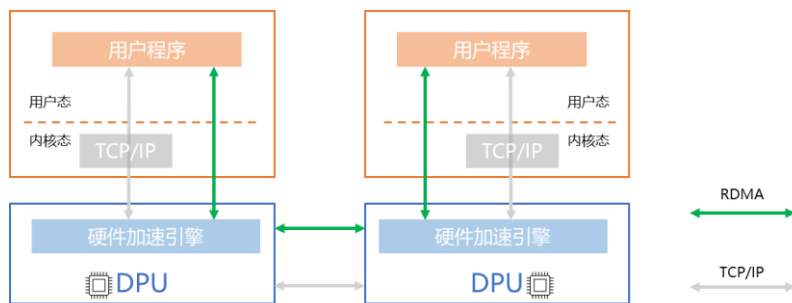


图 2-4 RDMA 协议与 TCP/IP 协议栈数据路径对比

使用 DPU 将 RDMA 协议（如 ROCEv2）卸载到硬件中，包括封装/解封装、拥塞控制算法等，并且支持单边操作，可以进一步降低被动端的 CPU 开销。主机安装 RDMA 协议驱动，向上对接 RDMA 原语，可以大大缩短时延并提升业务的吞吐率。

当前的 RDMA 技术的拥塞控制依赖 PFC，但是在开启 PFC 的应用场景下，PFC 反压会导致网络性能异常，尤其是在大规模组网的场景，会产生死锁等严重问题。为了避免全网 PFC 风暴，业界普遍在一个 Pod 内部署 RDMA 高速网络（千节点规模），云盘业务要求计算节点和存储节点在同一个 Pod 内，导致不同 Pod 间资源无法共享。在云场景下，网络规模较大，想要使用 RDMA 技术，核心在于对拥塞控制算法的优化，DPU 需要能够灵活进行算法的配置和编程，从而提供较好的动态时延。

2.3 存储系统卸载

在云服务中，DPU 可以为虚拟机或裸金属提供存储加速功能，通过软硬件结合方式实现存储协议卸载，灵活实现高存储 IOPS 性能和低主机 CPU 占用率的要求。

与虚拟化网络卸载原理类似，对于虚拟机或者裸金属，通过存储卸载可以提供灵活、高性能的弹性存储设备，存储系统卸载在 DPU 上完成的主要工作为：存储接口的卸载和存储网络协议栈加速。

2.3.1 存储接口卸载

在云计算场景中，DPU 作为云平台底座，可以为虚拟机或者裸金属实例提供标准存储接口 virtio-blk。利用 virtio 接口的前后端 vring 共享机制，为安装在 Guest OS 中的标准 virtio-blk 驱动提供存储数据交互，原生的 virtio-blk 接口已经可以实现动态的容量扩展和设备的热插拔。

NVMe 已经成为介质存储的行业标准，随着 NVMe 原生的设备在云数据中心出现，基于虚拟化的 NVMe 接口相对 virtio-blk 可以提供更高的带宽性能和更低的存储协议的延时。基于 NVMe 标准协议的虚拟化接口成为另一个标准，基于 NVMe 协议的存储设备可以支持更加丰富的功能和控制接口。

DPU 实现存储接口的卸载，主要方式为在 DPU 上实现存储接口后端，并提供 virtio-blk 或 NVMe 的设备接口给云主机，云主机中加载标准 virtio-blk 或 NVMe 驱动实现块存储的读写，无需额外的厂商专用驱动。

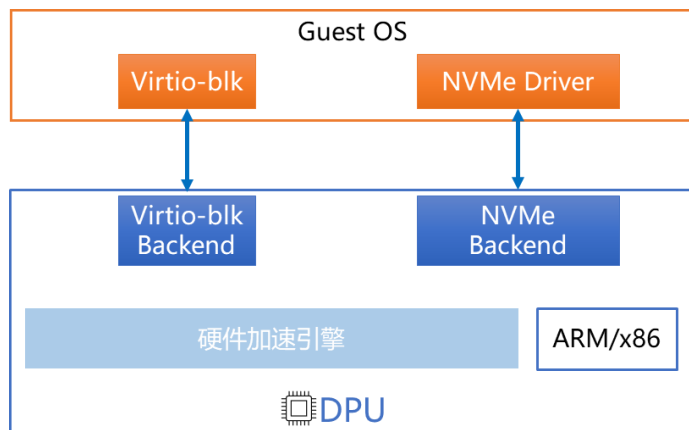


图 2-5 DPU 暴露 virtio-blk 或 NVMe 存储接口给云主机

2.3.2 存储网络协议栈卸载

DPU 对接收到的存储网络数据包基于存储网络协议栈进行相应的逻辑处理，当前业界多采用开源的存储协议比如 Ceph、iSCSI 或存储厂商自研的平台协议。将存储网络协议栈（或存储客户端）卸载到 DPU 上，可以减少主机侧 CPU 的开销，同时在裸金属场景下避免存储网络协议栈暴露给客户 OS 而引发的存储网络安全风险。

存储节点和运行在计算节点上的存储客户端一起工作提供了块存储级别的服务。在后端存储平台一般会支持不同的存储介质，如 HDD(SATA 和 SAS)、SATA SSD 和 NVMe SSD 等，从而构建出面向不同应用的存储服务。可以在 DPU 的 CPU 核中安装存储客户端处理 iSCSI 或 Ceph 的存储网络协议栈，配合 SPDK 用户态加速框架提升存储转发性能。

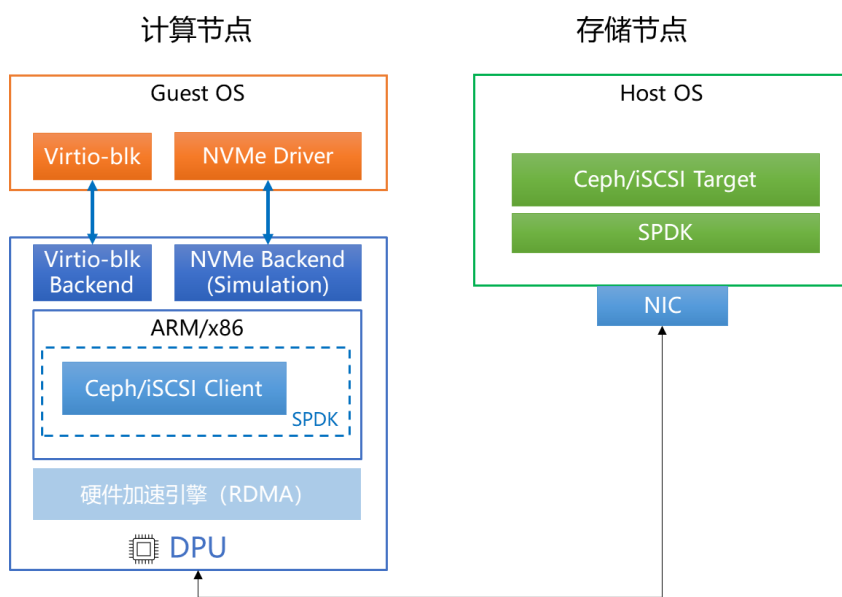


图 2-6 DPU 加速 Ceph、iSCSI 存储协议

更进一步，随着存储介质的发展和高性能存储需求的出现，存储接口正在向 NVMe 演进，而当前采用 iSCSI 或 Ceph 的存储网络协议栈需要针对 NVMe 的前端和存储后端进行协议转换才能实现标准 NVMe 设备，存储协议栈之间的转换导致用户无法使用到原生 NVMe 性能的存储设备。采用 NVMe-oF (NVMe over

Fabric) 协议可以进一步发挥 NVMe 高性能、低延迟和低协议负担的优势, 构建基于高速网络的 NVMe 共享存储系统互联结构。NVMe-oF 可以使用不同的高速的网络传输协议实现 NVMe 功能, 即可以支持多种类的 Fabric 网络, 包括 FC (Fibre Channel)、RDMA、TCP 等, 考虑到网络传输协议的高性能和兼容性问题, 业界选择 RDMA (RoCEv2) 居多。

在 DPU 上实现 NVMe-oF 协议栈的卸载加速, 可以在计算节点提供原生 NVMe 的后端存储接口, 通过高性能的 RDMA 网络协议 (如 RoCEv2) 连接到存储端, 在存储节点使用 DPU 硬件实现的 NVMe Target 管理 NVMe SSD, 整个存储网络传输端到端 bypass 主机 CPU, 并且没有任何的协议转换消耗, 为云主机提供了与本地 NVMe 性能接近的高性能弹性远端存储。

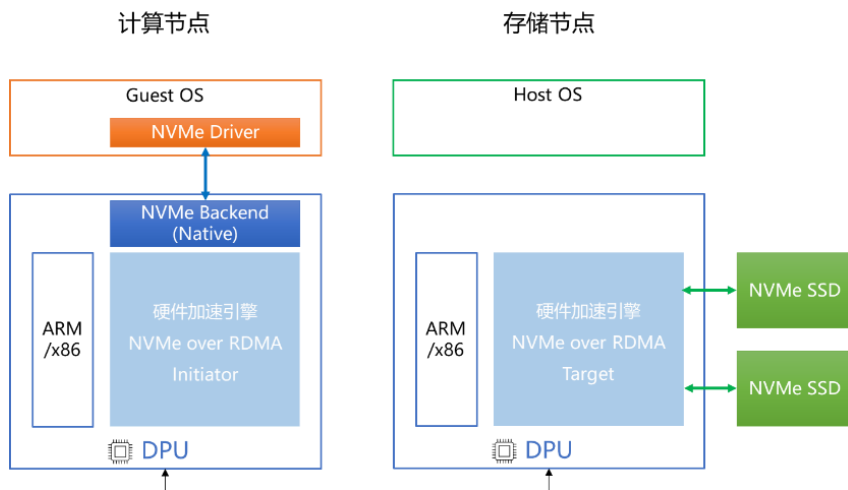


图 2-7 DPU 加速 NVMe-oF 存储协议

2.4 计算系统卸载

在虚拟化场景, 计算系统主要是指 KVM-QEMU 架构的 Hypervisor 系统, 虚拟层平台通过 Hypervisor 为虚拟化网元 (VNF, virtual network function) 提供虚拟机环境。计算系统卸载主要指的是 DPU 对 Hypervisor 的卸载。传统 Hypervisor 需要预留 10%~20% 的 CPU 计算资源, 通过将 Hypervisor 分成前后端架构, 前端运行在主机侧, 但只保留一些必要的内存标脏功能以及逻辑 CPU 的上下文同步等功能, 后端卸载到 DPU, 可实现 Hypervisor 预留计算资源降低到接近“零”, 并且可以减少 VM-Exit 带来的 VM 抖动对业务性能的影响。参考实现方式如下:

- 前端主要完成基本的虚拟机管理功能：
 - 与内核中 KVM 模块交互，完成对虚拟机的 vCPU 和内存的分配功能
 - 与后端的设备模型交互，确定虚拟机 IO 和内存的映射表
- 后端主要完成设备模型和生命周期管理功能：
 - 设备模型基于 QEMU 的 QOM 机制，完成虚拟设备的初始化，同时，在虚拟机启动或关闭时，与前端进行交互，完成内存和 vCPU 的资源分配和释放
 - 与 Libvirt 交互，完成虚拟机的生命周期管理功能

为了实现主机 CPU 资源的极致利用率，Host OS 需要尽可能地减少资源占用，由于大部分 IO 均由 DPU 直通到虚拟机，Host OS 基本上不需要进行 IO 硬件设备的管理，同时 Host OS 只需要支持运行虚拟机 vCPU 线程以及管理服务代理程序即可，因此可以考虑不使用基于 systemd 的复杂初始化系统以及 rpm 包管理工具，直接采用更为精简的初始化系统，实现主机 CPU 资源占用接近“零”。

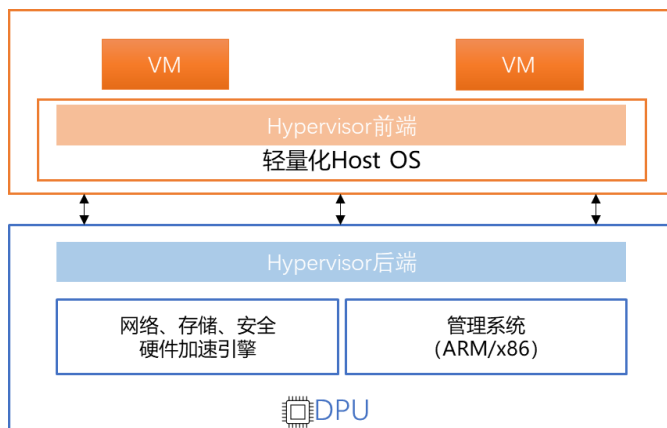


图 2-8 虚拟化计算系统卸载

2.5 安全系统卸载

云计算场景安全领域，主流虚拟化技术（KVM、XEN、VMware）中虚拟化漏洞广泛存在，Hypervisor 作为虚拟机的运行底座一旦存在漏洞将危及运行其上的所有虚拟机，甚至影响宿主机的安全。基于安全需求，各云厂商在宿主机上部署安全功能产品，而安全功能串行在网络转发中，产品的转发性能对整体网络的

吞吐量和时延有较大影响。伴随网络、存储卸载，将安全功能从主机上剥离，卸载到 DPU 上是必然趋势。安全系统的卸载可以实现虚拟化的安全隔离和网络数据加解密卸载。

2.5.1 虚拟化安全隔离

引入 DPU 之后，虚拟化管理平面、网络平面、存储平面卸载到 DPU 上，与主机完全隔离，如果黑客从内部通过虚拟机攻破了 Hypervisor，获取了主机操作系统的 root 权限，也只能破坏该计算节点，不会影响到其它主机。如果黑客通过外界攻击主机，必须先攻破 DPU 系统，这样相当于在主机前面又加上一层主机系统进行安全防御，并且可以在 DPU OS 之上定义多种安全策略：例如隔离控制、安全组、状态防火墙、包检测等，从而提升主机安全可信能力，而主机端无需安装任何代理软件，对主机性能无任何影响。

2.5.2 网络数据加解密

DPU 实现的网络数据加密是在数据进入到网络之前进行的，反之，加密的数据从网络进入 DPU，经 DPU 解密后再交给主机侧 CPU 处理。网络数据加密实现方式按照 OSI 7 层模型可分为数据链路层，网络/传输层及应用层三类。数据链路层加密/解密典型应用为 MACSec，MACSec 作为 IEEE 802.1AE 标准，可以单独使用也可以与 IPSec 或 TLS/SSL 结合使用。由于是数据链路层的安全实现，其对完整的数据帧进行加密处理，甚至可以做到对 LLDP、LACP、DHCP 及 ARP 等进行安全防护。网络/传输层加密/解密典型应用为 IPSec，利用 IPSec 可搭建 VPN 网络，实现在 VPN 网络上的数据安全传输。应用层加密/解密的典型应用为 TLS，TLS 被广泛应用于 HTTPS，现在互联网上几乎所有的网站均使用 HTTPS 提供页面访问服务。DPU 可为上述网络安全应用提供加速器单元，使其在应用与网络的边缘实现的加密/解密。

3. DPU 技术挑战和推进建议

DPU 作为一种软硬协同的虚拟化架构，相较仅做网络加速的智能网卡而言，一方面需要考虑与 CPU 中运行的虚拟化软件栈对接，另一方面由于 DPU 硬件的特殊设计导致卡与服务器存在更多的适配对接需求，因此 DPU 产业发展与硬件市场的配合度存在强关联。能否理解行业需求，拉通软硬件形成完整解决方案成为 DPU 发展与落地的关键。

当前 DPU 解决方案百花齐放，在下表总结的业界主流 DPU 方案中，由于硬件路线选择不同，DPU 卡的硬件设计存在较大差异，同时，各厂家均开始布局自有配套软件生态，并呈现出各自为营的态势。

表 1-1 业界主流 DPU 厂商当前软硬件生态

厂家	硬件路线	软件栈
NVIDIA	ASIC+CPU	DOCA
Intel	ASIC+CPU FPGA+CPU	IPDK
AWS	ASIC+CPU	自研
阿里云	FPGA+CPU	飞天 OS
华为	-	DPAK
中兴	FPGA/DSA+CPU	ZDOSA
芯启源	NP+CPU	-
大禹智芯	FPGA+SoC	DAYU RMOS
益思芯	FPGA/DSA+CPU	RDK
云豹	FPGA/ASIC+CPU	-

纵观处理器芯片几十年的发展历程，定制产品总是逐渐走向硬件形态的标准化和软件生态的通用化，最终形成专用与通用并存的格局。DPU 作为继 CPU、GPU 之后的第三颗数据中心核心处理器，基于自身虚拟化需求定制软硬耦合 DPU 产品固然能够最大化发挥 DPU 产品优势，但是随着更多的数据中心提出

DPU 需求，这种方式的技术门槛和引入成本过高，难以复制。而商业 DPU 产品由于各自发展软件栈生态，也导致软件与 DPU 适配对接的成本较高。同时，DPU 与服务器整机的生态也不可忽视。从物理关系上看，DPU 是挂载在服务器主 CPU 下的 PCIe 子设备，但从顶层云管逻辑来看，DPU 是服务器的管理员，是提供数据加速、存储加速、安全管控及云管的云底座。因此，不同于传统网卡、HBA 卡、GPU 等标准 PCIe 设备，主从关系的倒换使得 DPU 对传统服务器硬件设计提出了特殊要求，主要体现在硬件供电、散热适配、带外管理、整机开关机及异常处理策略等方面。为了解决 DPU 部件与服务器整机的适配问题，国内外头部互联网云厂商采用了硬件全定制策略，即网卡部件与服务器整机均由大云厂商内部研发团队制定完整的产品方案，并联合具备硬件产品设计能力的厂商进行产品研发及制造。对于其他行业终端客户，随着各家业务云化不断深入，裸金属场景需求逐步显现。同时国内外涌现多家 DPU 厂商，厂商希望网卡产品能够尽量多的适配主流厂商的通用服务器，另一方面，面对市场的不确定性，服务器厂商也有适配多家卡的需求。因此，面对即将爆发的生态及快速变化的云市场，硬件产品长迭代周期迫使硬件设计需具备超前性，当前亟需启动针对 DPU 与通用服务器整机在硬件层面的标准化推进工作，完善新技术生态。

市场需求决定了 DPU 的软硬件设计，而软硬件的开发、适配需要高昂的人力、财力成本投入，当前专用化的 DPU 路径不利于生态的发展，在百花齐放的行业现状下更需要考虑通用性和标准化，拉通需求方、虚拟化软件厂商、DPU 厂商与服务器厂商，形成端到端的闭环生态体系。

在中国移动算力网络云化基础设施建设过程中，出现了 DPU 的多种场景需求，例如：在网络云场景，需要网络系统的卸载，加速虚拟化网络提升核心网网元的转发性能；在公有云和 IT 云场景，需要管理系统的卸载，提供云化、弹性的裸金属管理系统；在智算场景，需要 RDMA 技术为 AI 训练和推理计算任务提供低时延网络通信能力。

随着需求越来越广泛，技术的演进对 DPU 的落地带来了诸多挑战，关键挑战即标准化程度对 DPU 技术易用性的影响。为推动 DPU 实际落地进度，提升产

业对接效率，减少重复适配性工作，中国移动从软件和硬件两个方面提出以下 DPU 标准化建议，希望通过抛砖引玉，启发产业思考，与业界共同探索 DPU 标准化方案，促进产业的良性发展。

3.1 软件系统标准化

与 CPU 和 GPU 这种先有硬件再有软件生态的发展背景不同，DPU 承载的虚拟化软件已经较为成熟，DPU 在此基础之上实现功能和性能的提升，因此对于业界云平台以及 DPU 厂商来说，DPU 能否规模部署取决于 DPU 配套软件是否“能用”且“好用”。

中国移动结合典型 DPU 应用场景需求，分析 DPU 软件部署方式，结合开源社区方案成熟情况，给出 DPU 云化技术架构管理、网络、存储、计算和安全五大系统软件标准化建议，并结合中国移动业务需求场景推动软件标准化阶段式发展和落地：

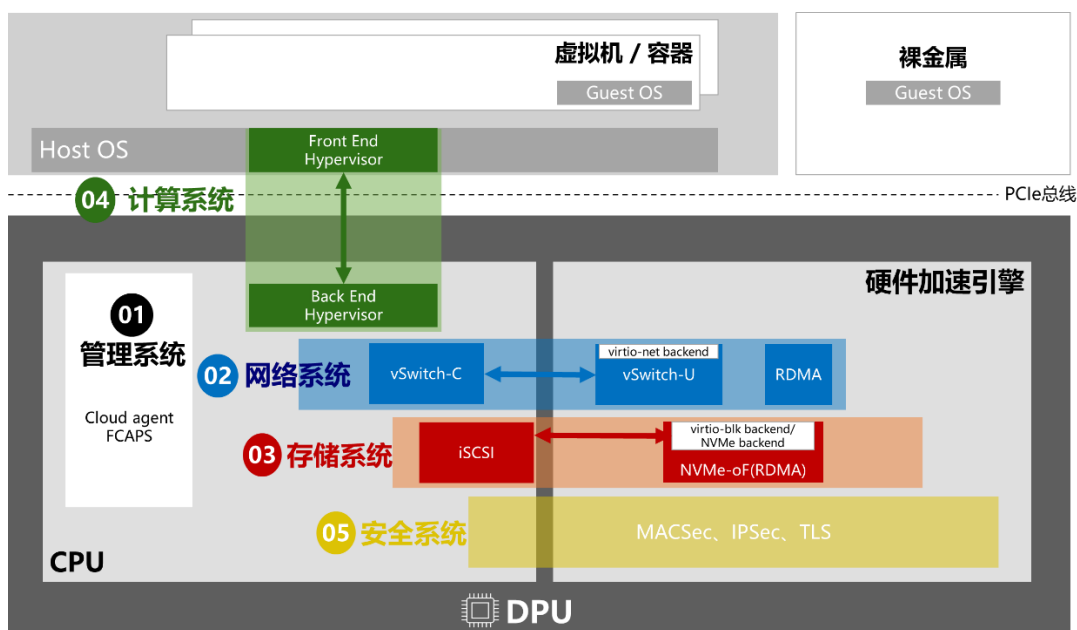


图 3-1 DPU 软件系统标准化

一、**管理系统**：引入 DPU 后，计算节点上的管理系统软件（OpenStack 等）安装位置从主机侧 CPU 调整到 DPU 的 ARM 或 x86 核上。管理系统软件主要实现虚拟硬件资源的逻辑管理功能，不使用 DPU 上的硬件加速引擎，因此理论上该部分软件与 DPU 天然解耦，DPU 需支持管理系统软件的编译、安装，可以通

过 DPU OS 和 Host OS 互通或者融合的方式，实现传统云平台管理软件在 DPU 上的无缝迁移部署和无感卸载，降低适配开发成本。虚拟化场景下，Libvirt 作为虚拟机管理组件也安装在 DPU 上，OS 资源视图发生变化，因此需要在主机侧 CPU 运行轻量化代理完成主机资源和虚拟机管理。此时需要保证 Libvirt 管理面接口和原生 Libvirt 接口保持一致，同时支持接口的扩展。裸金属场景下，可采用 Ironic 组件或增强 Nova 组件完成裸金属管理，DPU 厂商可以提供标准 Ironic Driver 集成到云管理软件中或适配标准 Nova 管理流程，实现裸金属生命周期的管理功能。

二、网络系统：虚拟交换网络 vSwitch 控制面运行在 DPU 的 CPU 核上，与 SDN 控制器对接，生成流表规则并通过 DPDK 标准 `rte_flow` 完成流表向转发面的下发，且需完成 `bond` 模式、`mtu` 值等数据通道参数的配置；转发面运行在 DPU 的硬件加速引擎上，负责流表匹配逻辑的硬件实现，并为虚拟机、容器或裸金属提供标准的 `virtio` 硬件后端。结合中国移动现网需求，提出以下标准化要求：

- 采用 DPDK 标准 `rte_flow` 接口完成流表从控制面到转发面的下发。同时考虑 NFV 业务网络拓扑需求，尤其在终端用户或基站通信时容易产生大规模流表瓶颈，影响转发性能，因此需定制流表收敛方案，规避大流表场景问题，同时降低业务倒换场景下流表下发通道阻塞的风险；
- 增强运维灵活性，业务网络平面数据全部在硬件加速引擎完成接收与发送，因此需要在数据在硬件加速引擎的生命周期流程中增加入口与出口的运维统计能力，便于故障分析定位，保证端到端可靠性；
- 网络接口采用用户态 vDPA 框架，在 DPU 上实现 `virtio` 后端，同一站点资源池内需统一要求 `virtio` 支持的 `feature`，以实现在部署不同厂商 DPU 的计算节点之间的虚拟机热迁移功能，保证云主机弹性特点。

RDMA 是超低时延网络的关键技术，但是由于传统 Socket 编程模型和 RDMA Verbs 编程模型差距显著，导致传统应用迁移到 RDMA 网络环境下需要进行大量代码改造，进而给业务带来不确定性风险或者可靠性风险，因此使用 RDMA Socket 适配层屏蔽底层 RDMA 原语，向上对接 Socket 接口成为必然趋

势，通过标准化的适配层，可以实现传统 Socket 应用无缝迁移到 RDMA 网络环境，在不改变业务代码的情况下，提高业务的响应时间和吞吐量。

三、存储系统：包含存储接口和存储网络协议两部分。基于开源 SPDK 框架提供存储卸载服务，利用 SPDK 的 JSON-RPC 配置接口，为虚拟机、裸机、容器提供一致的 virtio-blk/NVMe 虚拟块设备后端，基于 bdev 层抽象虚拟磁盘并提供 QoS、加密、压缩、DIF、EC、RAID 等高级特性。存储网络协议方面，iSCSI 协议运行在 DPU 上的 CPU 核上，需要支持标准 iSCSI-Initiator 的编译和安装；NVMe-oF 由 DPU 硬件加速引擎实现，需要实现 RoCEv2 等高性能 Fabric 网络。

四、计算系统：主要为 Hypervisor 计算虚拟化系统的前后端分离，在计算节点需要对 Hypervisor 软件进行合理的剪裁和卸载。主机侧保留轻量化的 Hypervisor 前端，KVM 完成 CPU 和内存的虚拟化管理，DPU 侧的 Hypervisor 后端配合 QEMU 完成虚拟机的初始化，并配合 Libvirt 完成虚拟机的生命周期管理。Hypervisor 前后端对接应标准化，同时需要对操作系统内核进行深度剪裁和修改，业界尚无参考模型，需要共同探索。

五、安全系统：DPU 可以提供节点出入口硬件级的数据加解密功能，通过虚拟化方式实现的分布式防火墙、DDoS 防御设备等安全应用可以利用 DPU 硬件级的数据处理能力，提升安全应用处理性能。此外针对 HTTPS 等应用层的加解密处理，需要提供统一的 API 接口来完成业务层用户密钥证书的管理。

针对以上五大系统在 DPU 上的卸载，OpenStack、DPDK、SPDK 等云计算虚拟化相关的主流开源技术框架已经逐步实现在 DPU 环境下的部署与应用，当前在裸金属管理和虚拟交换网络卸载两类功能需求下已经有较为成熟的技术方案。中国移动软件标准化推进将分为两个阶段进行：

- 第一阶段将聚焦于管理系统卸载、虚拟交换网络的卸载、iSCSI 存储网络协议栈卸载，主要解决核心网网络带宽升级导致的 CPU 处理瓶颈和满足弹性裸金属云化管理需求，利用开源技术方案，并增强平台端到端可靠性，制定技术方案，形成企业标准，实施现网试点验证；
- 第二阶段将聚焦于 RDMA 技术及接口标准化、NVMe-oF 高性能存储网

络协议卸载、Hypervisor 卸载、安全功能卸载等方面，满足 AI 训练场景低时延网络、大数据高性能存储读写、虚拟机“零”抖动、零信任安全等场景下更为极致的性能需求，这些领域在业界尚无规模性商用案例和成熟技术方案，需要共同探索和推进。

3.2 硬件系统标准化

当前 DPU 内除网络、存储卸载芯片外还会包含轻量级处理器芯片（或集成在其他芯片内的处理器内核）及带外管理芯片，该部件可看作在服务器整机内部的另一个小型服务器。在硬件标准化设计时，无论是对服务器自身还是承载的业务都应该遵循最小影响原则做通用增强，梳理绝对必要项，约束服务器整机能为 DPU 提供的最大能力。

从服务器整机角度，需重点考虑机构设计、供电散热和边带信号等方面的标准化工作：

- 机构设计：约束整机可支持卡的最大规格，避免异形卡及特殊机构件，减少机构适配工作量，减小应力可靠性风险；
- 供电散热：在槽位供电及辅助供电方面，约束整机可支持卡在整机 S5 状态启动，满足裸金属场景云底座管理诉求，同时高度关注 DPU 因业务需要不断增长的最大功耗，谨慎讨论并确定功耗门限；在配套散热能力方面，约束整机支持 S5 状态卡散热及调速；
- 边带信号：对于 PCIe 标准形态的 DPU 卡，在 PCIe 标准连接器定义基础上，约束整机能够提供给 DPU 的边带信号，落实信号类型、信号方向、电平、连接器选型、Pin 定义等细节，逐步引导行业做归一化设计。

从 DPU 角度，业务的多样性需求决定了当前硬件形态的多样化，行业内存在如 PCIe 卡、OCP 卡、自定义扣卡等多种形态，结合服务器整机的标准化方向，需要考虑以下方面：

- 关注服务器整机标准所能提供的机构外框上限及功耗上限；
- 避免主动散热器设计，并针对因高功耗引起的散热片材质升级进行必要

的力学前置仿真及后续测试；

- 预留必要的调试接口（通过线缆接入到 BMC 或者前挡板直出）。

由于 DPU 上存在 CPU 核，因此可以看作服务器整机的第二系统，这种架构的变化也导致服务器的带外管理方式产生了变化，从 DPU 与服务器整机带外管理交互角度，需要以下方面的标准化工作：

1. DPU BMC（或其他带外管理模块）具备的最小管理功能，例如：
 - 卡内信息抓取，远程 IPMI、Redfish 支持；
 - 散热信息提供至主板 BMC，基础散热信息应包括：进风口、出风口、主芯片；
 - 功耗监控、保活（看门狗）；
 - 故障诊断，主芯片故障上报；
 - 带外固件升级。
2. 裸金属场景服务器及 DPU 的上下电、复位、固件升级、异常下电、故障处理时的开关机流程及软件处理策略；
3. 云管业务对带外信息抓取方案，定义裸金属网卡 BMC 与主机 BMC 的核心职责，细化两者交互信息框架。

服务器整机硬件层面的标准化可让通用服务器与各厂家 DPU 具备适配前提，扫清新生态发展中的技术障碍。在这一过程中，需共同探讨 DPU 带外管理系统与服务器整机带外管理系统关系，明确 DPU BMC 应具备的核心功能，逐步收敛双 BMC 控制信息的交互框架，最终做到边界清晰。同时，中国移动倡议各行业终端客户关注因引入裸金属服务引发的组网方式变革，立足自身业务特征探讨网络业务平面、存储平面、带内管理平面、带外 BMC 管理平面的处理方式，明确 DPU 需提供的带内带外网口规格及数量，为 DPU 厂商提供产品路标。

4. 展望与产业倡议

为推动数字经济发展，我国陆续出台了新基建、东数西算等多项政策，加快构建以算力和网络为核心的新型基础设施。中国移动作为建设“数字中国、网络强国、智慧社会”的主力军，积极落实国家“东数西算”工程部署，高度重视算力布局，提出了算力网络全新发展计划。DPU 作为数据中心关键处理器，是解决数据中心云化需求、计算需求、网络需求与存储需求的关键枢纽。

中国移动面向网络云、移动云和 IT 云的建设，均考虑 DPU 技术的引入，当前，中国移动在 DPU 领域主要从两个方向进行推进：

- 标准开源方面，中国移动在 CCSA 完成 DPU 和智能网卡相关的标准立项，致力于从软件层面规范 DPU 的能力要求；在 ODCC 已经成立服务器硬件标准化项目，致力于研究 DPU 对服务器的需求以及裸金属服务器标准化方案。同时，中国移动也在紧密关注 OPI、DPDK、SPDK 等开源社区发展情况，秉承开放共赢的理念，与业界合作伙伴联合开展标准的讨论与制定。

- 试验落地方面，中国移动将依托算力网络试验网开展 DPU 技术方案验证，针对裸金属和虚拟化场景的管理、网络、存储等能力进行测试。在此欢迎更多的厂家参与 DPU 产品的测试验证，加速 DPU 产品的商用落地。

为促进 DPU 产业链良性发展，形成体系化的解决方案，中国移动希望发挥链长作用，以算力网络的多场景需求为引导，建立“DPU 开放实验室”，为 DPU 产业提供统一需求输出平台、方案孵化平台、生态合作平台。实验室将围绕探索应用场景、制定标准体系、构建产业生态三大工作内容，与合作伙伴共同推动 DPU 技术发展与产业繁荣。

附录：术语、定义和缩略语

缩略语	英文全名	中文解释
ARM	Advanced RISC Machines	RISC微处理器
ARP	Address Resolution Protocol	地址解析协议
ASIC	Application Specific Integrated Circuit	专用集成电路
BMC	Baseboard Management Controller	基板管理控制器
CPU	Central Processing Unit	中央处理器
DHCP	Dynamic Host Configuration Protocol	动态主机配置协议
DPDK	Data Plane Development Kit	数据平面开发套件
FC	Fibre Channel	光纤通道
FPGA	Field Programmable Gate Array	现场可编程逻辑门阵列
GPU	Graphics Processing Unit	图形处理器
HBA	Host bus adapter	主机总线适配器
Hypervisor	Virtual Machine Monitor	虚拟机监视程序
IaaS	Infrastructure as a Service	基础设施即服务
IB	InfiniBand	“无限带宽”技术
IO	Input/Output	输入/输出
IPSec	Internet Protocol Security	互联网安全协议
iSCSI	Internet Small Computer System Interface	Internet 小型计算机系统接口
KVM	Kernel-based Virtual Machine	基于内核的虚拟机技术
LACP	Link Aggregation Control Protocol	链路汇聚控制协议
Libvirt	用于管理虚拟化平台的开源的API	
LLDP	Link Layer Discovery Protocol	链路层发现协议
NP	Network Processor	网络处理器
NVMe	NVM Express	非易失性内存主机控制器接口规范
OvS	Open vSwitch	开源虚拟交换机
PCIe	Peripheral Component Interconnect express	高速串行计算机扩展总线标准
PFC	Priority-based Flow Control	基于优先级的流量控制
POSIX	Portable Operating System Interface of UNIX	可移植操作系统接口

QEMU	Quick EMUlator	虚拟操作系统模拟器
QOM	Qemu Object Module	QEMU对象模型
RDMA	Remote Direct Memory Access	远程直接数据存取
RPC	Remote Procedure Call	远程过程调用
RPM	Red-Hat Package Manager	红帽软件包管理器
S5	Soft Off	主板软掉电状态，部分器件如BMC、CPLD仍然工作
SATA	Serial Advanced Technology Attachment	串行高级技术附件
SDK	Software Development Kit	软件开发工具包
SLA	Service Level Agreement	服务级别协议
SPDK	Storage Performance Development Kit	存储性能开发套件
SR-IOV	Single Root I/O Virtualization	单根IO虚拟化
SSD	Solid State Disk	固态硬盘
SSL	Secure Sockets Layer	安全套接字协议
TCP	Transmission Control Protocol	传输控制协议
TLS	Transport Layer Security	安全传输层协议
vDPA	virtio Data Path Acceleration	Virtio数据路径加速
VNC	Virtual Network Console	虚拟网络控制台
VPN	Virtual Private Network	虚拟专用网络