

网络功能虚拟化的关键技术

Key Technologies in Network Function Virtualization

王路/WANG Lu
赵鹏/ZHAO Peng
付乔/FU Qiao

(中国移动通信有限公司研究院, 北京
100053)
(China Mobile Research Institution, Beijing
100053, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 02-0012-005

摘要: 指出了网络功能虚拟化(NFV)的5个关键技术: 硬件及管理技术、虚拟层技术、管理编排技术、可靠性技术、加速技术, 并介绍了NFV发展情况及当前存在的问题。认为NFV作为运营商网络转型的核心技术架构, 是虚拟化和云计算等信息技术(IT)技术在电信领域的一次大规模应用。随着技术的成熟, 未来将很快看到NFV架构的电信网络; 以NFV为出发点, 通信技术(CT)和IT将走向深度融合。

关键词: NFV; 管理和编排(MANO); Hypervisor

Abstract: In this paper, 5 key technologies of network function virtualization (NFV) are pointed out, including hardware and management technology, virtual layer technology, management arrangement technology, reliability technology, acceleration technology. And the development status and problems are also introduced. NFV is considered as the core technology architecture of the network transformation of operators, and it is the large-scale application of information technology (IT) technology (such as virtualization and cloud computing) in the field of telecommunications. With the maturity of the technology, the NFV architecture of the telecommunications network will be used in the future, and communication technology (CT) and IT will deeply integrate.

Keywords: NFV; management and orchestrator (MANO); Hypervisor

1 NFV 架构和应用场景

1.1 NFV 架构

网络功能虚拟化(NFV)的基础架构由欧洲电信标准化协会(ETSI)NFV行业规范组织(ISG)设计完成, NFV逻辑架构如图1所示。

NFV逻辑架构主要分为4个部分: NFV的管理和编排(MANO)系统用于整体编排和控制管理; NFV基础设施(NFVI)提供网元部署所依赖的基础设施环境; 虚拟网络功能(VNF)这一层包括虚拟网元自身以及负责管理VNF的网元管理系统(EMS); 运营支持系统/业务支持系统(OSS/BSS)是运营支撑系统。

1.2 NFV 应用场景

NFV的应用范围非常广泛, 从网络边缘到网络核心, 从固定网络到移动网络, 所有网络功能的实现都有可能重新设计或改造。以下介绍几种NFV应用的典型场景[1]。

(1) 固定接入网。

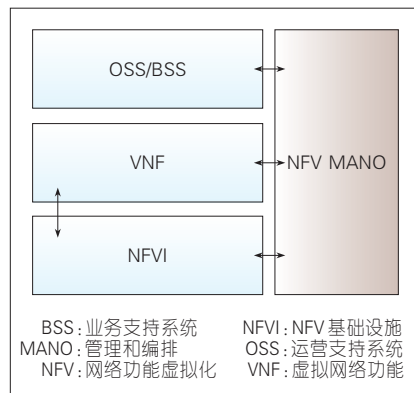
虚拟客户终端设备(vCPE)和虚

拟宽带远程接入服务器(vBRAS)是NFV部署的典型示例之一。vCPE将复杂的网络功能及新增业务以虚拟化方式部署在网络侧而非用户侧, 同时为运营商未来新增业务乃至第三方业务提供开放平台。vCPE的部署使得传统固定接入网络变得更加灵活, 用户可以根据需求定制自己的网络。从目前的产业发展现状来看, 固定接入网络的NFV化已经成为业界的共识, 但目前技术方案尚未完全成熟, 各类VNF的功能仍有待完善。

(2) 移动核心网。

核心网虚拟化一直是NFV应用的重点领域。目前, 全球运营商在该领域概念验证(POC)、试点乃至部署案例众多, 比如: 日本DoCoMo公司尝

试自主集成构建物联网解决访问虚拟演进的数据核心网(vEPC)。同时, 中国移动也以虚拟IP多媒体子系统(vIMS)和基于IMS的语音业务(VoLTE)为切入点, 进行了深入的



▲ 图1 NFV逻辑架构

收稿日期: 2017-01-20
网络出版日期: 2017-03-07

NFV 试点。

(3) 移动接入网。

移动接入网的分布式特征所带来的高成本、高管理复杂度的挑战使得 NFV 成为其未来发展的重要解决方案。新一代基站布建架构——云化无线接入网(C-RAN)就是一种典型应用场景。C-RAN 通过将基带处理器从站点移开,并置于核心更深处,可有效降低设备成本,改善协作,增加网络容量。同时,在无线接入网络方面,基于虚拟技术的无线控制器(AC)虚拟化和池化技术也逐渐引起广泛关注。AC 虚拟池的实现可以有效降低 AC 设备成本,增强设备的可靠性,提高 AC 设备利用率,简化运营商运维管理 AC 设备的复杂度。

(4) 数据中心应用。

虚拟防火墙(vFW)、虚拟负载均衡器(vLB)和虚拟安全套接层(vSSL)/Internet 协议安全性(IPSEC)网关(GW)当前在数据中心也得到了大量应用。得益于 NFV 带来的灵活性特点,上述虚拟化网元可以灵活扩容和缩容。其次,基于 SDN 的业务链功能,能够实现非常灵活的增值业务编排,真正实现用户按需定制。

2 NFV 关键技术及其解决方案

2.1 硬件及硬件管理技术

2.1.1 硬件概述

硬件作为 NFV 最底层的基础设施,涉及的硬件包括 3 类:计算类、存储类和网络类。

(1) 计算。

计算类硬件采用商用通用服务器(COTS),目前业界主流的指令集为 X86,服务器为双中央处理器(CPU)的 2 路服务器。主流服务器形态包括 5 种:刀片式、机架式、多节点(也称为高密度服务器)、整机柜、机柜级架构(RSA)。

- 刀片式服务器是一种刀框集

成多个卡式服务器单元(形态像刀片)的服务器,刀框内同时集成背板、交换背板、管理单板、电源和风扇等部件。其集成度较高,可以节省大量机柜空间。

- 机架式服务器是一种集成 CPU、内存、主板、网卡、硬盘、电源和风扇的服务器。一个机架式服务器为一个计算单元。

- 多节点服务器是一种在特定空间的框内,集成电源、风扇和多个服务器单元的服务器,多个服务器单元共享电源和风扇。

- 整机柜服务器是一种标准化的服务器,整机柜内集成电源模块、风扇模块、交换模块、管理模块、服务器模块。

- RSA 是 Intel 提出的一种未来服务器形态,是芯片资源池化,通过资源池虚拟出需要的服务器的技术。

(2) 存储。

当前 NFV 业界主流存储技术采用 IP 存储区域网络(IP-SAN)和分布式存储,相应的存储介质为磁阵和存储型服务器。

磁阵是一种可靠性达 5 个 9、性能高的成熟存储硬件。存储型服务器是一种配置硬盘多,满足大容量存储需求的服务器。不同存储介质需要结合相应的存储技术,以发挥最大的技术优势。

(3) 网络。

网络类硬件主要采用交换机。为满足站点组网的要求,实现设备间互通,采用接入交换机和核心交换机实现站点组网。接入交换机一般采用 1U 的盒式交换机,主要实现二层互通,核心交换机采用模块交换机,主要实现三层互通。

2.1.2 硬件资源池共享

NFV 业务种类繁多,主要集中在无线接入、固网接入、核心网和部分业务平台等。硬件资源池的共享是实现资源共享的第一步,网元部署地理位置和业务特性,决定了硬件资源

有无共享的必要,硬件配置则决定硬件资源池有无共享的可行性。因此,在相同地理位置,网元业务特性相同的网元,应保证其硬件配置的一致,同一个硬件资源池中应尽量减少硬件配置的种类。

2.1.3 硬件管理

硬件管理是 NFV 网络部署、运营、保证业务质量必不可少的部分。然而当前硬件管理规范方面并不是十分标准。

服务器一般采用简单网络管理协议(SNMP)和智能平台管理接口(IPMI)实现管理,不同厂商 SNMP 和 IPMI 存在差异。磁阵一般采用 SNMP 和存储管理接口标准(SMI-S)实现管理,SNMP 方面不同厂商存在较大差异,SMI-S 为全球存储网络工业协会(SINA)组织主导的国际标准,是统一的标准;然而 SMI-S,在故障告警方面没有涉及。交换机可采用 SNMP 实现管理,不同厂商在实现方面存在较大差异。

目前台式系统管理任务组(DMTF)正在制订服务器、磁阵和交换机的相关技术标准,即 Redfish 和 Swordfish。主流服务器、磁阵厂商均在推动该标准的成熟。Redfish 和 Swordfish 均基于 Restful 应用程序编程接口(API)实现,接口符合未来发展的趋势。Redfish 和 Swordfish 在故障管理方面暂未成熟,DMTF 正在大力推动其完善

2.2 虚拟层技术

虚拟层是基础架构层 NFVI 的重要组成部分,包含 Hypervisor 和虚拟化基础设施管理(VIM)。虚拟层将物理计算、存储、网络资源通过虚拟化技术转换为虚拟的计算、存储、网络资源池,提供给虚拟机使用,同时,提供虚拟资源的管理和运维^[1]。

2.2.1 VIM

虚拟化基础设施管理平台,负责

对虚拟化资源进行统一的管理、监测控制、优化,对虚拟机进行生命周期管理等。

目前,主流 VIM 平台基于 OpenStack 社区进行开发,包括身份认证及授权、虚拟机镜像管理、计算资源管理、存储资源管理、网络资源管理、虚拟机生命周期管理等能力。同时,VIM 平台整合了 NFVI 的运维和管理能力,包括虚拟化层的关键绩效指标(KPI)监测控制,故障告警收集及上报等。

2.2.2 虚拟化软件 Hypervisor

Hypervisor 将通用物理服务器与上层软件应用分开,使得具有不同操作系统的多个虚拟机可以在同一个物理服务器上运行,最大化地利用硬件资源,即一个物理服务器的硬件资源可以被多个虚拟机共享。Hypervisor 把硬件相关的 CPU、内存、硬盘、网络资源全面虚拟化,并提供给上层 VNF 使用,具备计算虚拟化、存储虚拟化和网络虚拟化能力。Hypervisor 可以与 VIM 系统交互实现对虚拟机的创建、删除等操作以及故障管理、性能管理等功能。

(1) 计算虚拟化。

实现对服务器物理资源的抽象,将 CPU、内存、输入/输出(I/O)等服务器物理资源转化为一组可统一管理、调度和分配的逻辑资源,并基于这些逻辑资源在单个物理服务器上构建多个同时运行、相互隔离的虚拟机执行环境,实现更高的资源利用率,同时满足应用更加灵活的资源动态分配需求。

为实现 5 个 9 的高可用性,NFV 对计算虚拟化提出了新的技术要求,包括 CPU 核绑定、大页内存、非均匀存储器存取(NUMA)、亲和性/反亲和性和部署等能力,同时需禁止 CPU、内存的超分配。

(2) 存储虚拟化。

将存储设备进行抽象,以逻辑资源的方式呈现,统一提供全面的存储

服务。可以在不同的存储形态,设备类型之间提供统一的功能。

(3) 网络虚拟化。

在服务器的 CPU 中实现完整的虚拟交换的功能,虚拟机的虚拟网卡对应虚拟交换的一个虚拟端口,服务器的物理网卡作为虚拟交换的上行端口。

目前,业界主流的网络虚拟化技术包括虚拟交换机及单根 I/O 设备虚拟化技术(SR-IOV)。虚拟交换机(OVS)是在开源的 Apache2.0 许可下的产品级质量的虚拟交换标准,通过虚拟化软件提供的部署在主机上的虚拟交换功能,主机节点的物理接口及虚拟机的虚拟网卡(vNIC)分别与虚拟交换机连接,通过虚拟交换机实现与外部网络的数据传输。SR-IOV 基于虚拟功能的虚拟机直通,虚拟机直接使用物理网卡资源进行网络通信,减少传统的虚拟交换带来的 CPU 消耗,提升性能,减少时延。

2.3 管理编排技术

2.3.1 管理编排

NFV MANO 的架构由 NFV 编排器(NFVO)、VNF 管理器(VNFM)、VIM 组成,完成对于 NFV 系统内虚拟资源、虚拟网元和网络服务的管理。MANO 系统架构图 2 所示。

NFVO 实现网络服务和网元管理及处理,提供网络服务生命周期的管理。VNFM 实现虚拟化网元 VNF 的生命周期管理,包括 VNF 实例的初始化、VNF 的扩容/缩容、VNF 实例的终止。VIM 是虚拟化基础设施管理系统,主要负责虚拟基础设施的管理,监测控制和故障上报,面向上层 VNFM 和 NFVO 提供虚拟化资源池。VIM 提供虚拟机镜像管理功能。

2.3.2 管理编排产业发展

MANO 由 ETSI NFV ISG 首先提出,并于 2014 年底发布 MANO 阶段 1 规范,明确了 MANO 系统架构、功能

实体、接口和参考流程,为业界 NFV 管理系统的设计提供了参考。

目前 NFV MANO 相关标准化工作尚未完成接口和模版的数据模型定义,尚无法指导各厂家设备基于统一格式开发并实现互通;与此同时,很多开源社区提供了开源版本的 MANO 或 MANO 部分组件,形成了对标准的重要补充。

在 NFVO 层面,目前最重要的开源组织是刚刚宣布成立的开放网络自动化平台(ONAP)组织,ONAP 由中国移动主导的 OPEN-O 和 AT&T 主导的 ECOMP 合并成立。在 VIM 层面,OpenStack 已经成为 VIM 的事实标准,多数厂家 VIM 基于 OpenStack 实现,并支持 OpenStack API,供 VNFM 和 NFVO 调用。

2.3.3 NFV 后的网络管理

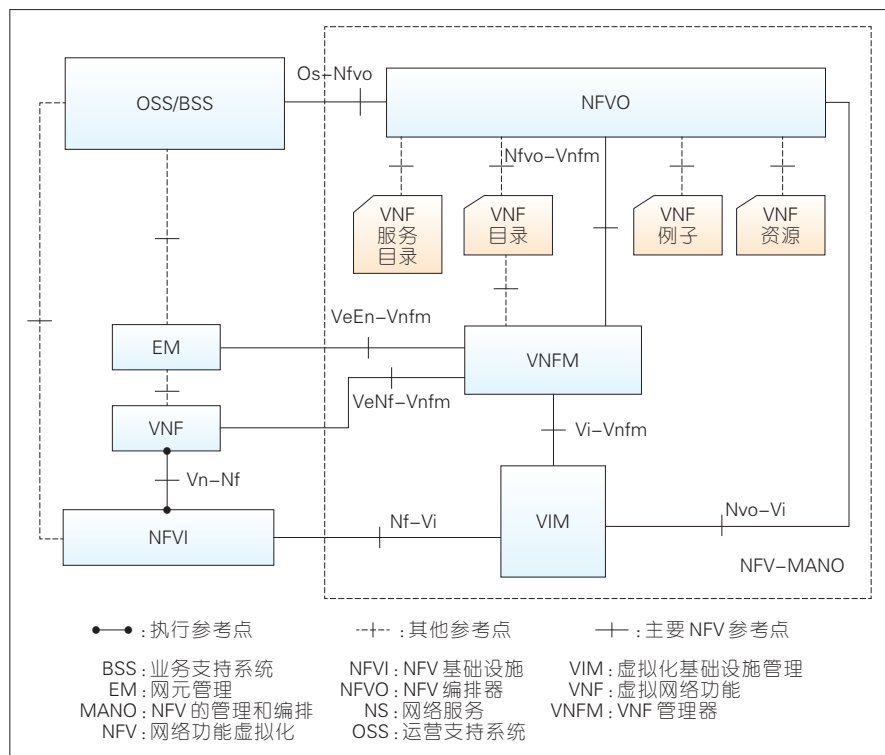
MANO 的引入,实现了网络的灵活管理和动态调整,同时也为运营商运维带来了全新的挑战。

首先,MANO 引入了 NFVO、VNFM、VIM 等一系列新的管理实体,将原有烟囱式的一体化运维分拆成了资源和应用两个层面,对现有 OSS 网管管理架构和流程均造成较大影响,需着重考虑 MANO 与 OSS 的协作关系。

其次,NFV 引入了 NFVI、VNF、网络服务(NS)等一系列新的管理对象,对现有 OSS 管理的资源管理模型、配置模型、性能模型、故障模型和多层故障关联等均造成较大改变。

再次,NFV MANO 向运营商提供了更为智能化的网络管理手段,有望使网络运维由传统的故障驱动型运维转向依靠预定义的策略和模版,依靠大数据和机器学习,实现网络的自动调整和治愈。这对未来网管及网络运维人员的技能均提出了极大的挑战。

基于以上挑战,未来网管的一种形态将发生大的变化,在管理模式上分为设计态和运行态。设计态采用



▲图2 MANO系统架构

形式化语言提供系统运行的策略和编排的模版,供机器执行,实现网络的自动调整和治愈;运行态监测控制资源和应用,对其进行实时的调度和治愈。

2.4 可靠性技术

引入虚拟化技术后,如何保证虚拟网元依然能够表现出与传统物理设备相当的可靠性成为人们关心的主要问题。虚拟软件和云管理技术可靠性要求较低。如何基于相对不可靠的虚拟化和云技术提供高可靠的电信业务呢?

NFV的可靠性可以自底向上,分别从硬件、虚拟云平台、虚拟网元3个层次实现。

(1) 硬件可靠性。

硬件层面的可靠性即包括NFV所在的硬件节点的可靠性,也包括物理网络、存储的可靠性。通过多年信息技术(IT)和通信技术(CT)的积累,这些设备的可靠性已经有了较为完备的部署方案,基本可以满足NFV

的需求。

(2) 虚拟云平台可靠性。

虚拟云平台的可靠性包括云管理平台的可靠性和虚拟管理平台(Hypervisor)的可靠性。目前OpenStack已被普遍认可作为云管理平台。在IT和CT关于OpenStack可靠性的需求中,流传着一个“牲口还是宠物”的玩笑。IT领域认为虚拟资源就是“牲口”,需要成群的管理,却不需要对每一个都格外关注,当某一个虚拟资源出现故障后,我们只需要为用户重新启动一个虚拟资源补充这个空白即可。然而,在CT领域,由于其虚拟资源上运行的是电信网元,每一个故障都可能导致电信业务的错误乃至瘫痪,因此CT运营商必须把这些虚拟资源看成是“宠物”,即对每一个虚拟资源都要格外关注,一旦出现故障,就要及时发现并恢复,以保障其上业务的正常运行。

Hypervisor的可靠性问题目前业界未能达成一致意见。虽然存在一些私有的Hypervisor的可靠性解决方

案,然而广泛采用的开源的Hypervisor并没有可靠性保障,而主要是依靠云管理节点发现故障并进行修复。这就导致Hypervisor的可靠性机制完全依赖于云管理平台及其二者之间网络的可靠性。同时,云管理节点目前发现故障及修复的接口尚未完善,因此在Hypervisor发现故障和修复方面存在较大问题。

(3) 虚拟网元可靠性。

传统的电信网元软件一般都有较为完善的可靠性机制。当引入虚拟化技术后,由于引入了虚拟层,电信网元软件无法直接读取到网元硬件的信息,而只能看到其依赖的虚拟层信息。因此,引入虚拟化后,电信网元的可靠性方案也要因此进行相应的修改和优化,以满足电信网络的快速故障发现和恢复要求。目前,大部分虚拟网元在交付同时都能提供可靠性方案。然而各种可靠性方案存在差别,且对硬件和云平台的要求各异。

2.5 数据面加速技术

传统的IT通用服务器采用的多核处理器的包处理性能无法满足通信网络数据面网元的高性能要求,因此出现了多种数据面加速技术。传统支撑包处理的主流硬件平台大致可分为3个方向:硬件加速器、网络处理器、多核处理器。

2.5.1 硬件加速器

硬件加速器由于本身规模化的固化功能具有高性能、低成本的特点。专用集成电路(ASIC)和现场可编程门阵列(FPGA)是最广为采用的器件。

ASIC是一种应特定用户要求和特定电子系统的需要而设计、制造的集成电路。ASIC的优点是面向特定用户的需求,在批量生产时与通用集成电路相比体积更小,功耗更低,可靠性提高,性能提高,保密性增强,成本降低等;但是ASIC的灵活性和扩

展性不够,开发费用高,开发周期长。

FPGA 作为 ASIC 领域中的一种半定制电路而出现,与 ASIC 的区别是用户不需要介入芯片的布局布线和工艺问题,而且可以随时间改变其逻辑功能,使用灵活。FPGA 以并行运算为主,其开发相对于传统 PC、单片机的开发有很大不同,以硬件描述语言来实现。相比于 PC 或单片机的顺序操作有很大区别。

2.5.2 网络处理器

网络处理器(NPU)是专门为处理数据包而设计的可编程通用处理器,采用多内核并行处理结构,提供了包处理逻辑软件可编程的能力,在获得灵活性的同时兼顾了高性能的硬件包处理。其常被应用于通信领域的各种任务,比如包处理、协议分析、路由查找、声音/数据的汇聚、防火墙、服务质量(QoS)等。其通用性表现在执行逻辑由运行时加载的软件决定,用户使用专用指令集即微码进行开发。其硬件体系结构大多采用高速的接口技术和总线规范,具有较高的 I/O 能力,使得包处理能力得到很大提升。

NPU 拥有高性能和高可编程性等诸多优点,但其成本和特定领域的特性限制了它的市场规模。

2.5.3 多核处理器

多核处理器在更为复杂的高层包处理上具有优势,随着包处理开源生态系统逐渐丰富,为软件定义的包处理提供了快速迭代的平台。现代 CPU 性能的扩展主要通过多核的方式进行演进。这样利用通用处理器同样可以在一定程度上并行地处理网络负载。由于多核处理器在逻辑负载复杂的协议及应用层面上的处理优势,以及越来越强劲的数据面的支持能力,它在多种业务领域得到广泛的采用。再加上多年来围绕 CPU 已经建立起的大量成熟软件生态,多核处理器发展的活力和热度也是其

他形态很难比拟的。

当前的多核处理器也正在走向片上系统(SoC)化,针对网络的 SoC 往往集成内存控制器、网络控制器,甚至是一些硬件加速处理引擎。

多核处理器集成多个 CPU 核以及众多加速单元和网络接口,组成了一个 SoC。在这些 SoC 上,对于可固化的处理交由加速单元完成,而对于灵活的业务逻辑则由众多的通用处理器完成,这种方式有效地融合了软硬件各自的优势。

2.5.4 VNF 加速

对于性能要求一般的控制面/数据面网元,可以直接部署在通用多核处理器服务器上执行。

对于性能要求严苛的数据面网元,可以考虑采用辅助硬件加速器的方式。取决于业务功能的定制化和灵活性考虑,综合成本因素,选择 ASIC 或者 FPGA 加速器。

3 结束语

NFV 作为运营商网络转型的核心技术架构,是虚拟化和云计算等 IT 技术在电信领域的一次大规模应用。目前以 ETSI NFV 架构为技术架构,运营商和业界厂商大力推动 NFV 的分层解耦和资源池化,并且在固定接入网、移动接入网、移动核心网、数据中心等场景下开展试验验证和商用尝试。

虽然 ETSI 在 NFV 架构上已经定型,但在具体模块、接口、流程等实现上还不完善,目前业界的开源社区、标准组织和厂家乃至运营商都在积极推动相关技术的进步和成熟。文章中,我们详细分析了采用 NFV 分层解耦后之后,需要关注的关键技术。首先需要通过虚拟化技术在硬件资源池之上形成虚拟资源池,并且考虑硬件共享和硬件管理、虚拟资源管理的问题,完成虚拟资源生命周期管理;其次,通过管理和编排系统对各类资源形成视图,完成虚拟网元的生

命周期管理和网络服务的管理,并且解决各层故障上报和故障关联的问题,同时还要处理好 NFV 管理编排和 OSS 的关系,形成新的网管体系。电信业务有着高可靠性和实时性等要求,因此仅仅实现 NFV 分层解耦无法满足这类特殊要求。要实现电信业务的 5 个 9 可靠性要求,需要从下向上在每个层面都提供可靠性保障,并且各层面能够进行联动,提供系统级别的可靠性;要实现电信业务实时性的要求,数据面网元功能需从硬件实现到系统设计都进行针对性的加速。

以 NFV 为基础的运营商网络转型大幕已经开启,随着技术的成熟,未来将很快看到 NFV 架构的电信网络,以 NFV 为出发点 CT 和 IT 将走向深度融合。

参考文献

- [1] 李正茂. 通信 4.0[M]. 北京,中信出版社, 2016
- [2] Network Functions Virtualization (NFV); Management Orchestration: GS NFV-MAN 001 V0.6.2 (2014-07)[S]. ETSI, 2014

作者简介



王路,中国移动通信研究院网络技术研究所项目经理、高级工程师;主要从事核心网、NFV 相关的研究工作。



赵鹏,中国移动通信研究院网络技术研究所研究员;主要从事 NFV/SDN、核心网、OSS/BSS 相关的技术研究和标准化工作。



付乔,中国移动通信研究院网络技术研究所项目经理、中级工程师;主要从事 NFV 相关的研究和开源工作。