



专题: SDN/NFV

基于 SDN/NFV 的未来网络实验平台

魏亮¹, 李蓉¹, 吴正东¹, 张玉亮¹, 程智炜¹, 黄韬²

(1. 江苏省未来网络创新研究院, 江苏 南京 211100; 2. 北京邮电大学, 北京 100876)

摘要: 依据 SDN/NFV 集中管控、动态、灵活、高效、可编排等特点, 提出了基于 SDN/NFV 技术的未来网络实验平台的构建方案。该平台主要采用 OpenStack 和 OpenDaylight 的开源架构, 同时研发设计 SDN 跨域虚拟网络通信、虚拟网元管理以及网络服务编排三大关键技术, 实现了底层异构资源的实时动态管理与开放共享。同时, 根据实验用户对网络资源的需求, 灵活按需编排各种网络资源与 SFC 服务, 为用户提供端到端的网络实验验证服务。

关键词: 软件定义网络; 网络功能虚拟化; 虚拟网元; 网络服务编排

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017097

Future network experiment platform based on SDN/NFV

WEI Liang¹, LI Rong¹, WU Zhengdong¹, ZHANG Yuliang¹, CHENG Zhiwei¹, HUANG Tao²

1. Jiangsu Future Networks Innovation Institute, Nanjing 211100, China

2. Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract: According to the characteristics of SDN/NFV, the future network experiment platform based on SDN/NFV was proposed. The open source architecture of OpenStack and OpenDaylight was used, and the SDN cross-domain communication system, virtual network element management and network service orchestration were researched. Real-time dynamic management and effective sharing of the underlying heterogeneous resources were achieved. At the same time, according to the needs of the experimental users of the network resources, various network resources and SFC service was flexibly orchestrated and end-to-end network experimental verification services were provided to users.

Key words: SDN, NFV, virtual network element, network service orchestration

1 引言

随着互联网技术和传统行业的深度融合, 应用服务种类与需求愈发多样化, 尤其对网络服务

质量的需求不断上升, 基于 TCP/IP 的分布式传统网络架构面临着越来越严峻的挑战。为此, 研究人员提出了许多新型的算法、协议和网络架构。然而, 新算法和新协议从提出到实际应用往往需

收稿日期: 2017-03-10; 修回日期: 2017-04-01

基金项目: 国家高技术研究发展计划(“863”计划)基金资助项目(No.2015AA016101, No.2015AA016105)

Foundation Items: The National High Technology Research and Development Program (863 Program) (No.2015AA016101, No.2015AA016105)



要长期的实验验证,为了获得可靠的实验数据,需要通过建设专门的网络实验平台为这些创新性的技术提供真实、独立的实验环境。

近年来,软件定义网络(SDN)引发了广泛的关注,基于SDN的通用网络实验床也得到了广泛的建设,如GENI-OpenFlow^[1]、RISE^[2]等。SDN采用了将网络控制平面与转发平面分离新架构,将控制平面统一到独立的控制器中,并且能够通过开放可编程接口,动态改变网络状态和结构^[3,4]。这种动态、灵活可编程的特性为研究人员验证新算法、新协议以及网络架构创新提供了便利的手段。另外,为了满足网络实验对网络多样、多变的需求。SDN实验床需要具备更灵活、更开放、更简单的网络虚拟化能力,要求能够动态调用底层物理资源,按需为实验者分配虚拟实验资源,并将这些资源与服务灵活自动编排后提供给实验者使用。

针对上述情况,本文在利用SDN技术的基础上引入了NFV、网络服务编排等创新技术,设计开发了基于SDN/NFV的未来网络实验平台。平台首次采用了将“业务模式与技术模式整合到一起”的理念,将包括SDN/NFV、云计算技术、5G、物联网等在内的新一代网络需求都整合到架构中,以更快的速度、更灵活的方式为实验者按需构建差异化的网络实验环境。

2 基于SDN/NFV的未来网络实验平台

基于SDN/NFV的未来网络实验平台,基于OpenStack和OpenDaylight实现计算、存储、网络资源的虚拟化,在物理资源共享复用的基础上保证不同网络实验环境之间相互隔离,互不干扰;实现多个跨域实验节点无缝融合,支持跨域虚拟网络的端到端通信;能够根据实验需求定制虚拟网元(virtual network element)、构建网络拓扑以及配置链路带宽,并且根据业务场景需求动态编排实验资源;具有实验管理、实验拓扑管理、镜

像管理、基础设施管理等功能,同时采用自主研发的SDN跨域虚拟网络通信、虚拟网元管理以及网络服务编排三大关键技术,能够很好地满足多样化网络创新实验需求;支持SDN、NFV、5G、物联网、网络安全等多种创新实验类型,并根据实验场景,按需为实验者构建弹性可扩展且相对独立的网络实验环境。其整体架构如图1所示。

其中,SDN跨域虚拟网络通信系统能有效地兼容底层的异构基础设施,实现虚拟网络资源的动态管理和跨域通信;虚拟网元管理系统通过定义各种虚拟网络元素,提供高度自定义的虚拟网络拓扑环境的能力,用户可以根据实验对网络的需求,按需构建面向应用的网络资源;服务编排系统通过引入业务编排及服务链功能,将各类网络服务抽象出来,再组合运用服务链技术为用户提供端到端一体化的服务。

2.1 基于SDN的跨域虚拟网络通信

在建设网络实验平台的时候,往往单一数据中心资源无法满足用户需求。多节点数据中心跨域虚拟网络通信就成为急需解决的问题。基于SDN的跨域虚拟网络通信系统,不仅能灵活地构建跨域多节点数据中心的通信链路,其开放可编程性更能满足各种复杂的跨域业务需求。

本文采用的跨域虚拟网络通信系统架构如图2所示,主要包括支持VxLAN协议的跨域网关设备、SDN控制器和跨域虚拟网络通信系统等部分。使用SDN控制器作为跨域网关节点的管理中心实现域间网络自动化的编排管理。在跨域虚拟网络通信系统中根据实验用户的虚拟网络需求对数据中心间的资源进行动态调度,使用VxLAN^[5] overlay的域间链路满足实验平台灵活组建跨域网络的要求,通过VNI映射管理实现跨域虚拟网络的互联互通,通过OpenFlow协议下发流表来保证域间链路的QoS。

本文提出的跨域通信系统,其核心实现了包括租户VNI(VxLAN ID)映射,租户创建的虚拟机MAC地址或IP地址域所在的资源节点IP地址

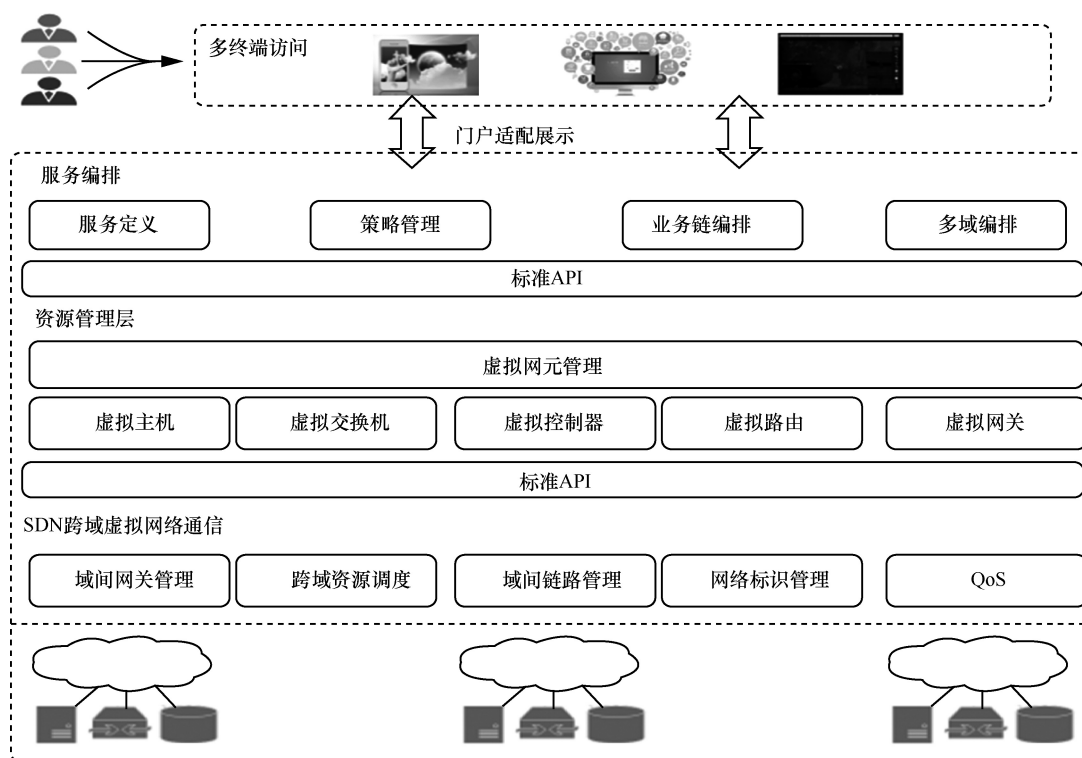


图1 未来网络实验平台整体架构

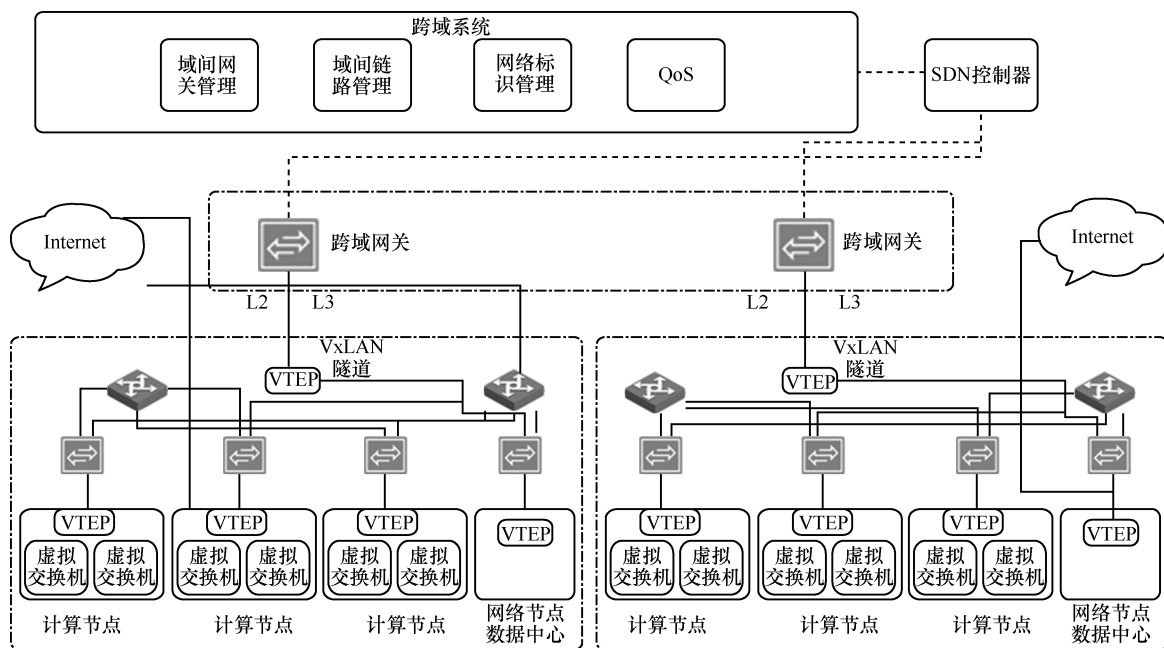


图2 基于 SDN 的跨域虚拟网络通信系统结构

映射以及不同虚拟网络的域间 QoS 保证。具体的实现原理如图 3 所示。

2.1.1 租户网络 VNI 标识转换

实验用户可以在多个数据中心域创建资源切

片，系统为每一个租户网络分配一个跨域 VNI，作为域间通信的标识。图 3 中所示用户创建实验时，实验平台同时请求分配位于数据中心 A 的虚拟网元 A1 和数据中心 B 的虚拟网元 B1。两个虚

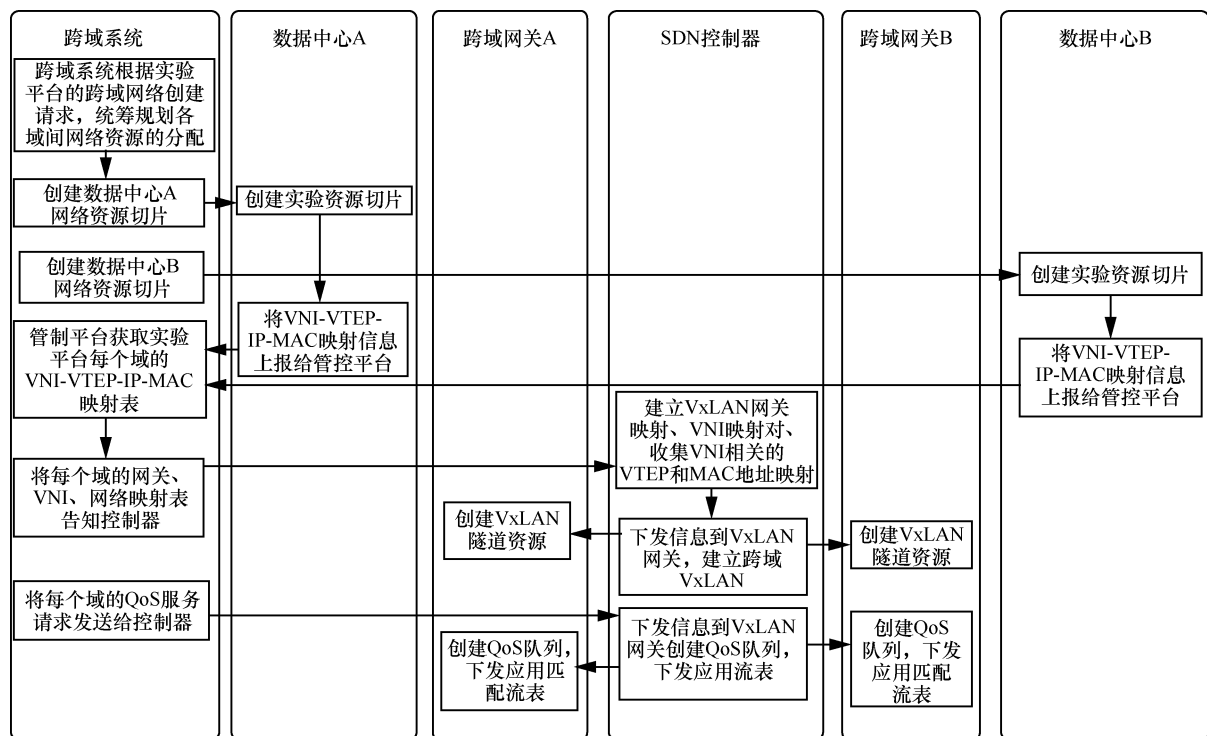


图3 基于SDN的跨域虚拟网络通信系统原理

拟网元所在的租户网络拥有不同的VNI标识。跨域虚拟网络通信系统能够管理不同租户网络的VNI标识,不同OpenStack域间通信时由跨域SDN控制器给跨域网关下发VNI映射流表实现VNI的转换,由一边OpenStack域的虚拟网络VNI标识转换成另一边OpenStack域的虚拟网络VNI标识。这样就可以实现跨越两个OpenStack域的虚拟网络内各网元的互联通信。

2.1.2 虚拟网元地址信息与计算节点的映射

实验用户创建的虚拟网元都是在数据中心计算节点中,虚拟网元A1发送数据给虚拟网元B1的过程是虚拟网元A将数据分组发送给本身的虚拟网卡,在通过虚拟网卡将数据分组发送给虚拟网元A1所在的计算节点,再从计算节点发送到跨域网关;虚拟网元B1所在的数据中心对应的跨域网关处理转换VNI之后,需要知道该数据分组发送给数据中心B中的哪一个计算节点。因此需要管理用户创建的虚拟网元的MAC地址、IP地址和所在计算节点的映射关系

来规划转发路径,并通过跨域SDN控制器给跨域网关层下发流表实现跨域虚拟网元间的端到端通信。

2.1.3 基于队列的域间链路QoS保证

实验平台需要根据实验用户的需求来灵活定义域间链路QoS参数。通过跨域SDN控制器配置网关设备间的队列,不同的队列享用的带宽或优先级不同,以进行拥塞管理。SDN控制器能够响应packet-in消息,分析数据分组的业务类型,动态下发流表到跨域网关输出到不同的队列中。对于域间通信服务质量要求高的实验服务,跨域虚拟网络通信系统能够根据域间通信服务质量等级标识,请求SDN控制器给跨域网关下发带宽保证QoS队列以及相应的高优先级QoS匹配流表。

2.2 虚拟网元管理

当前网络架构正随着SDN/NFV等网络创新技术的引入发生着根本性变革,导致学术科研领域迫切需要能够支持新架构、新设备、新系统的网络实验系统。直接基于物理设施搭建网络实验

环境, 不仅功能固化无法满足创新实验要求, 而且需要投入大量资金。同时由于网络环境与物理硬件紧耦合, 无法实现复杂的网络场景和高效的资源复用。实验平台中借鉴 SDN/NFV^[6]相关理念提出虚拟网元的概念, 对应现实的物理网元, 虚拟网元不是具体的网络设备, 而是实现特定网络功能的虚拟网络元素。虚拟网元不仅包含实现了各种网络功能的虚拟网络设备, 还包含虚拟设备间进行通信所必需的虚拟网络链路。这样实验平台就拥有在有限的基础设施基础上构建高度自定义的虚拟网络拓扑环境的能力, 从而可以适应不同网络实验的需求。

虚拟网络设备根据设备实现的网络功能不同可分为虚拟控制器、虚拟交换机、虚拟主机和虚拟网关等。通常使用虚拟机来构建虚拟网络设备, 当然还可以采用其他方式来构建, 如使用 Docker 容器等。在构建设备之前需要对设备实现的指定网络功能以及需要的资源要求、软件条件进行规划和描述。如构建一个包含 8 个端口的虚拟交换机和一个虚拟主机相连的拓扑, 可以使用如图 4 所示的数据格式进行定义。

```
{
  "topology": {
    "name": "single example",
    "nodes": [
      {
        "id": 1,
        "name": "simple switch",
        "function": "switch",
        "os": "ubuntu",
        "cpu": 2,
        "ram": 512,
        "disk": 5,
        "ports": 8
      },
      {
        "id": 2,
        "name": "simple server",
        "function": "compute",
        "os": "centos",
        "cpu": 4,
        "ram": 4096,
        "disk": 50,
        "ports": 1
      }
    ],
    "links": [
      {
        "source_node": 1,
        "source_port": 8,
        "target_node": 2,
        "target_port": 1
      }
    ]
  }
}
```

图 4 虚拟网元描述定义示例

虚拟网络链路泛指为了实现虚拟网络设备之间进行通信而采用的技术, 相当于实现现实中网线的功能, 可以保证逻辑连接的虚拟网络设备间进行正常的通信。在构建虚拟网络链路时可选择性地指定链路的各个指标特征, 如带宽、时延、QoS 等。实验网络中虚拟网络链路根据链路中流量类型可分为数据平面链路、控制平面链路。数据平面链路只进行数据的交换不涉及控制管理, 而控制平面用于虚拟控制器对虚拟交换机之间的管理。

下面简单介绍实验网络中虚拟网络设备和虚拟网络链路到 OpenStack 虚拟环境的映射转化, 由虚拟机实现虚拟主机 (vHost)、虚拟交换机 (vSwitch)、虚拟控制器 (vController) 等虚拟网络设备, 由网络 (network)、子网 (subnet) 实现虚拟链路 (vLink) 保证虚拟网络设备间的连通。具体实验网络与 OpenStack 虚拟环境对应关系如图 5 所示。

根据 OpenStack 网络对象的实现原理, 网络是二层连通的, 通过将虚拟主机和虚拟交换机接入同一网络对象中, 则可以实现二层通信。在实验平台中选择以 VxLAN 隧道模式来构建网络, 虚拟机之间的连接关系如图 6 所示。

在系统实现过程中, 如图 7 所示。在虚拟交换机上连接一个虚拟网关 (vGateway), 虚拟网关通过一条虚拟链路 (vLink) 与外部网络互连, 可以实现虚拟主机访问外网。在 OpenStack 中需要将虚拟网关与路由 (router) 接入同一子网对象中, 子网提供 DHCP 服务, 为虚拟网关分配 IP4 地址, 路由配置网关地址 IP3, 虚拟网关与路由实现三层连通; 并且路由配置外网地址 IP1 连入外部网络, 通过静态路由实现内部网络与外部网络互通; 对应于实验网络中, 虚拟网关即可与外部网络实现通信, 其连接关系如图 7 所示。类似的可以通过将虚拟控制器和虚拟交换机放入同一子网对象中, 为虚拟控制器与虚拟交换机分配同一

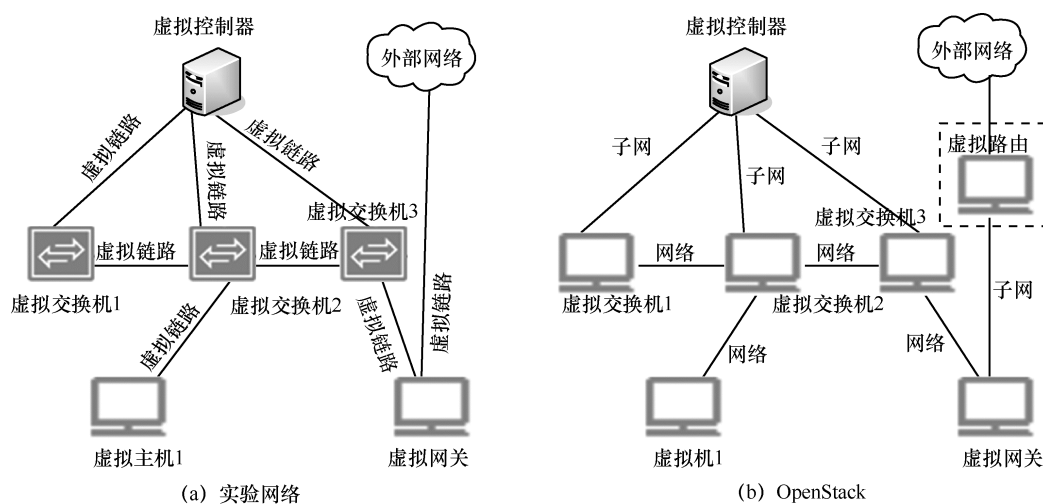


图5 实验网络与 OpenStack 虚拟环境对应关系

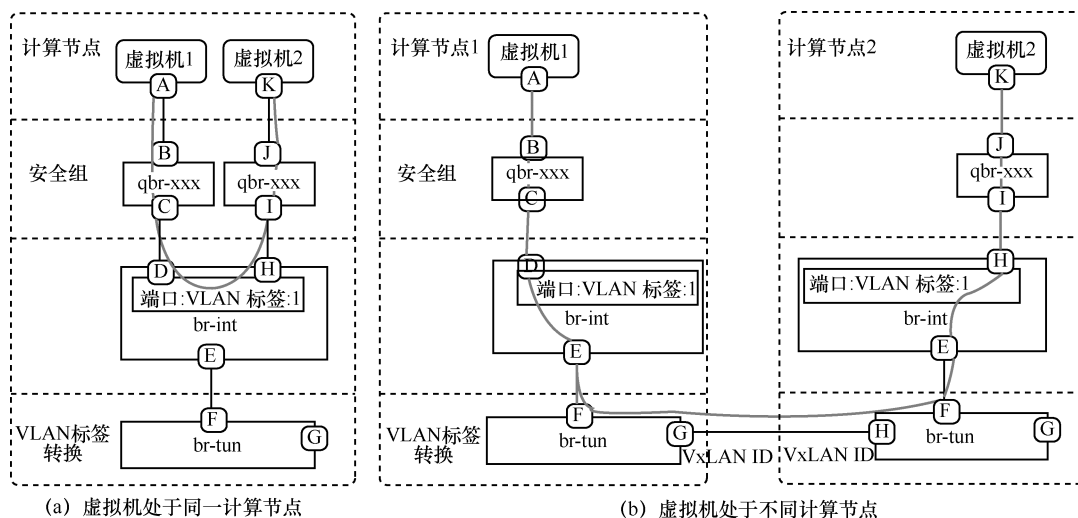


图6 同一网络中虚拟机间二层通信（其中灰线表示虚拟机之间的通信路径）

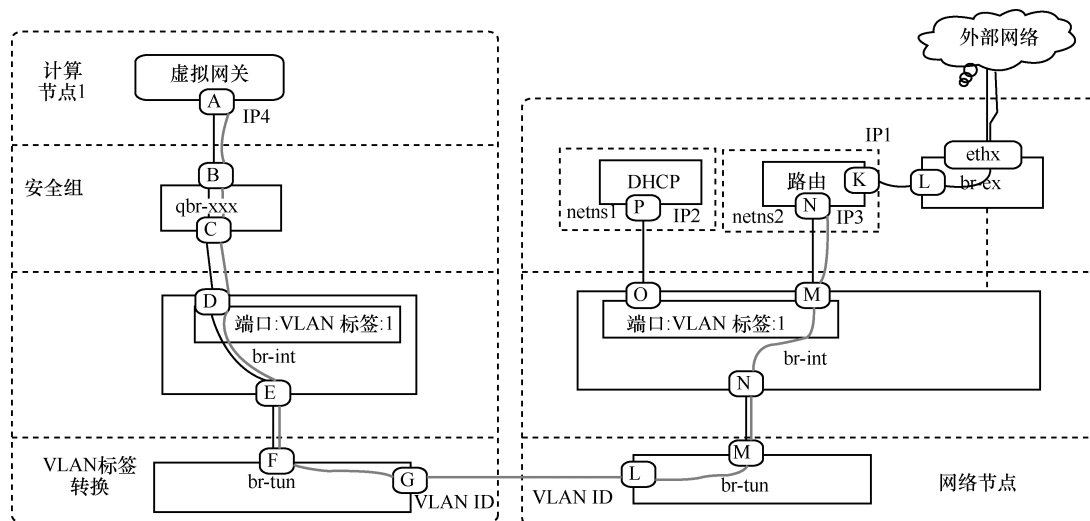


图7 虚拟网关访问外部网络（其中灰线表示虚拟机之间的通信路径）

网段的 IP 地址, 实现三层连通, 从而实现虚拟控制器对虚拟交换机的管控。

2.3 网络服务编排

在管理编排网络服务中, 尤其是在面对多租户的环境下, 网络业务复杂, 数据报文在传递过程中需要经过多层业务节点来保证提供安全、稳定、快速的服务。其中典型的业务节点有防火墙、NAT、入侵防御系统 (IPS) 以及其他应用场景下的网络服务 (如 vIMS、vCPE) 等。通过业务逻辑定义网络流量按照一定顺序通过这些业务点, 即所谓的服务链。随着 SDN/NFV 技术不断成熟, 服务链的重要性不言而喻。通过引入业务编排及服务链功能, 网络服务可以被抽象出来, 再组合纳入服务链及服务图的概念中, 为用户提供端到端一体化的服务。下文将分别从技术架构及相关案例来阐述本平台的网络服务编排功能。

本平台通过 SFC (service function chaining, 服务功能链) 技术实现服务链功能, 通过 NSH^[7] (network service header) 为负载流加上特定的网络服务头部, 数据转发设备通过识别该头部特征进行数据流转发, 由此将服务拓扑与实际网络拓扑分离^[8]。服务链中的各个角色: 业务节点 (service function)、分类节点 (classifier)、转发节点 (service function forwarder)、SFC 代理 (SFC proxy) 协同完成业务流定义与编排。流进入转发域后被分类节点映射至对应的服务链, 转发节点接受业务流并根据服务链描述将其转发至业务节点, 若业务节点无法处理 NSH, 则该流会被导入 SFC 代理节点完成 NSH 操作, 并沿定义路径进行转发^[9]。实验平台使用 OVS (Open vSwitch) 作为数据平面转发节点, 使用 VPP (vector packet processing) 作为分类及代理节点。

实验平台中通过网络控制器对虚拟化网络进行集中控制, 控制器可以根据业务需求生成服务

链流规则。不仅支持虚拟环境下网络功能实例间业务流编排, 也支持将物理环境下的硬件设备 (如硬件防火墙、DPI 等设备) 接入虚拟网络成为业务节点提供网络服务, 提升了实验平台的可扩展性及网络业务的灵活性。

为了满足实验用户在网络业务动态编排、虚拟网元和物理设备联动实验等方面的需求, 实验平台通过 NFVO 统一编排业务服务链、VIM 域以及网络控制器, 通过不同层次的实验平台 API、VIM-API、控制器 API 实现精细化的资源调度, 利用 SFC 技术实现服务链的动态编排。图 8 展示了实验平台在多域、虚实结合的部署环境下服务链的实现方案。

在图 8 示例中, 通过实验平台 NFVO 组件提供的接口完成服务链组成节点及业务流特征定义, 申请及注册 VNF1、VNF2、PNF、VNF3 成为服务链中的业务节点, 使 VNF1 与 VNF3 之间通信流量经过 VNF2、PNF, 接收请求后 NFVO 完成业务节点与虚拟资源映射, 并通过 VIM 提供的接口对各 VIM 域资源完成相关资源申请与调度。VIM 接收到 NFVO 的请求后, 调用网络控制器接口对相应节点及其所在域进行 SFC 相关配置, 生成分类节点并下发流规则。VIM 接收 SFC 实例化请求后流程如图 9 所示。

3 结束语

基于 SDN/NFV 的未来网络实验平台, 主张为实验建网, 构建面向实验场景、感知智慧资源的能力。采用自主研发的基于 SDN 的跨域虚拟网络通信系统, 实现了跨域资源的实时动态调度和有效共享利用; 通过虚拟网元管理系统, 提供构建高度自定义网络功能和网络拓扑的能力; 此外, 采用网络服务编排技术为实验编排各种网络资源与网络服务, 实现资源随业务而动的端到端一体化服务。

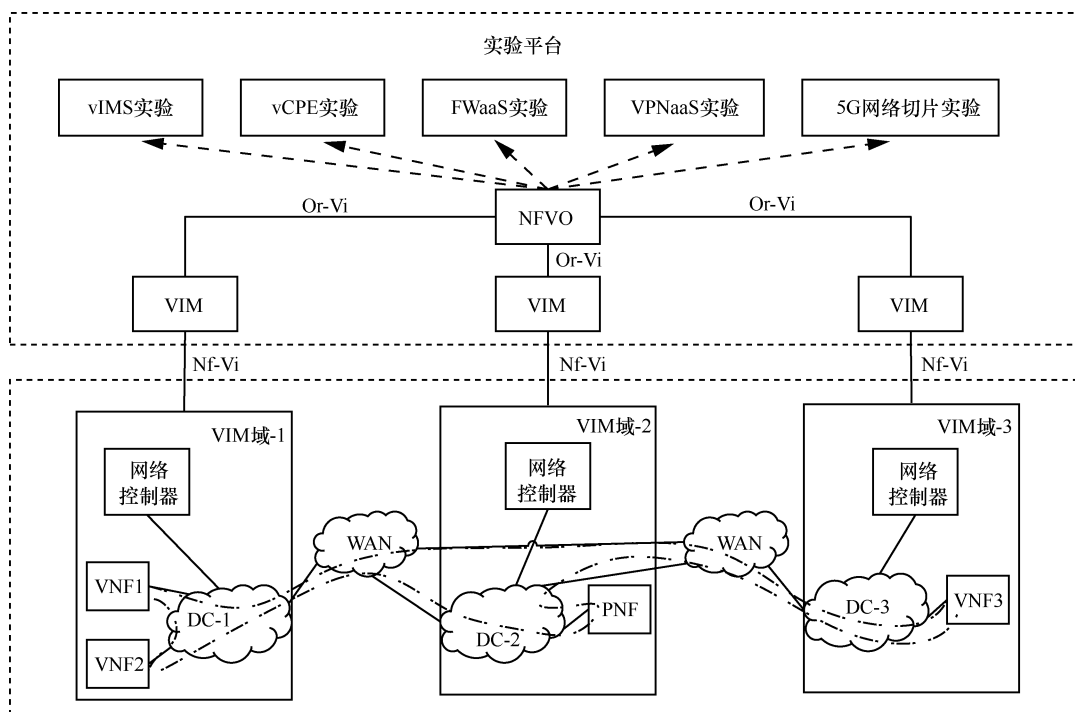


图8 多域 VIM 及虚实结合案例

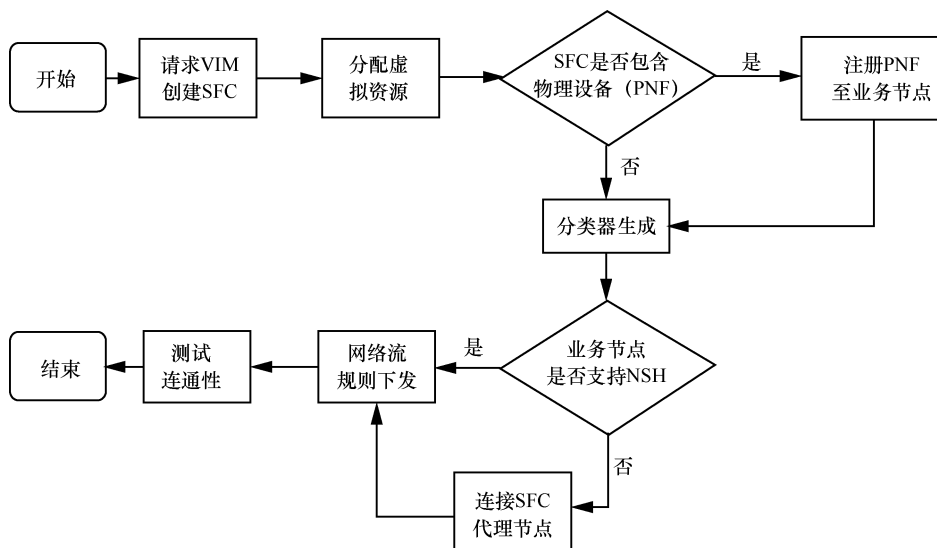


图9 实验平台实例化 SFC 流程

参考文献:

- [1] GENI. GENI OpenFlow[EB/OL]. [2017-01-20]. <http://groups.geni.net/geni/wiki/OpenFlow>.
- [2] KANAUMI Y, SAITO S, KAWAI E, et al. RISE: a wide-area hybrid OpenFlow network testbed[J]. IEICE Transactions on Communications, 2013, 96(1): 108-118.
- [3] 黄韬, 刘江, 魏亮, 等. 软件定义网络核心原理与应用实践[M]. 北京: 人民邮电出版社, 2014.
- [4] Open Networking Foundation. Software-defined networking: the new norm for networks[EB/OL]. [2017-01-20]. <https://www.opennetworking.org/>.
- [5] HODDA S K, KAPADIA S, KRISHNAN P. Using Trill, FabricPath, and VXLAN: designing massively scalable data centers (MSDC) with overlays[M]. Indianapolis: Cisco Press, 2013.

- [6] OASIS. Topology and orchestration specification for cloud applications: version 1.0. 25[S]. 2013.
- [7] The Internet Engineering Task Force (IETF). Service Function Chaining (SFC) Working Group (WG)[S/OL]. [2017-01-25]. <https://datatracker.ietf.org/wg/sfc/charter/>.
- [8] MIJUMBI R, SERRAT J, GORRICO J, et al. Network function virtualization: state-of-the-art and research challenges[J]. IEEE Communications Surveys & Tutorials, 2016, 18(1): 236-262.
- [9] SONKOLY B, CZENTYE J, SZABO R, et al. Multi-domain service orchestration over networks and clouds: a unified approach[C]//SIGCOMM 2015, August 17-21, 2015, London, UK. New York: ACM Press, 2015: 377-378.

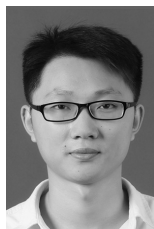
[作者简介]



魏亮 (1981-), 男, 江苏省未来网络创新研究院 CDN 团队总监, 主要研究方向为软件定义网络、云计算、网络试验平台。



李蓉 (1981-), 女, 江苏省未来网络创新研究院高级工程师, 主要研究方向为软件定义网络、网络虚拟化、云计算。



吴正东 (1991-), 男, 江苏省未来网络创新研究院工程师, 主要研究方向为软件定义网络、网络虚拟化、云计算。



张玉亮 (1988-), 男, 江苏省未来网络创新研究院高级工程师, 主要研究方向为软件定义网络、网络虚拟化、云计算。



程智炜 (1993-), 男, 江苏省未来网络创新研究院工程师, 主要研究方向为软件定义网络、网络功能虚拟化、云计算。



黄韬 (1980-), 男, 北京邮电大学教授, 主要研究方向为未来网络体系架构、内容中心网络、软件定义网络。