



NFV 业务链资源分配技术

唐 宏, 罗雨佳

(中国电信股份有限公司广州研究院 广州 510630)

摘 要: SDN/NFV 为运营商实现网络重构和架构演进提供了支撑技术, 业务链及其组网技术是 NFV 的一个重要应用场景, 相关的业务链资源分配问题是业界的研究热点。首先对业务链现状进行了梳理, 重点分析了已有的资源分配机制和存在的问题, 提出了一种业务链资源分配方案, 充分考虑了对网络资源的优化利用, 实现了 VNF 实例放置和业务链路径优化选择。实验结果表明, 本方案具有更大的系统用户容量和更短的业务路径。

关键词: 业务链; NFV; 资源分配; 路径选择

doi: 10.11959/j.issn.1000-0801.2015300

NFV Service Chain Resource Allocation Technology

Tang Hong, Luo Yujia

(Guangzhou Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China)

Abstract: SDN/NFV provides supporting technologies on network re-architecture and evolution for network operator. Service chain and related networking technology is becoming a significant use case of NFV, of which resource allocation is one of the key research points. The service chain status in quo was studied. The resource allocation mechanisms and related problems were analyzed. After that, a service function chain resource allocation scheme, which optimizes the usage of the network resource and realizes virtual network function instance deployment and route selection, was proposed. Experiment results show that the proposed method is able to serve more traffic.

Key words: service chain, NFV, resource allocation, route selection

1 引言

NFV(network function virtualization)技术基于通用服务器硬件,并以软件化的形式提供各类网络功能,为运营商未来网络业务的灵活部署和网络结构的快速调整提供了极大的支撑能力,尤其是相关的业务链(service chain)技术,可以灵活地将固移网络中的网元和网络功能,如 DPI(deep packet inspection,深度分组检测)、服务器负载均衡器、NAT(网络地址转换)等,以 VNF(virtualized network

function,虚拟化网络功能)的形式按需组合,快速形成各类强大的网络业务处理引擎。

NFV 业务链是一个有序的 VNF 集合,用户流量按照指定的策略依次通过多个 VNF,例如不同套餐的流量通过不同的 VNF 集合,以实现网络业务的按需处理。由于业务链中的 VNF 运行在通用服务器上的虚拟机(VM)环境中,各服务器之间的 VNF 又组成不同的虚拟化网络,这种高度可伸缩的电信业务处理架构保证了 VNF 的灵活加载和快速调整能力,但也引入了复杂的资源调度问题。

收稿日期:2015-09-01;修回日期:2015-11-03

论文引用格式:唐宏,罗雨佳. NFV 业务链资源分配技术. 电信科学, 2015300

Tang H, Luo Y J. NFV service chain resource allocation technology. Telecommunications Science, 2015300



目前,运营商网络业务链应用中的突出问题^[1]表现在业务链描述模型的建立、业务链资源分配机制、虚拟化组网、业务链自动配置与管理等几个方面。其中,业务链资源分配机制直接影响了 NFV 系统运行和资源集约利用效率,具有重要的研究意义。

本文将针对业务链在运营商网络的应用场景,分析相关的资源分配问题,并提供一个 NFV 场景下的业务链资源分配方案。

2 NFV 业务链应用场景及业界研究现状

2.1 NFV 业务链在 IP 网络中的应用场景

NFV 业务链可应用于宽带城域/骨干网边缘、IP RAN、移动核心网等领域。根据目前的 NFV 服务器处理能力和网卡容量,应用于 IP 网络边缘或作为移动域 P-I 网关较为合理。

目前,IP 网络边缘的专用硬件存在着功能较少、新业务板卡开发周期长、部署缺乏灵活性等问题,通过在 x86 服务器上部署软件化的 VNF 及其相关的业务链部件,可以降低部署成本和加快部署周期。如图 1 所示,可以在 IP 网络边缘 BRAS(broadband remote access server,宽带远程接入服务器)的位置部署各类基于 x86 服务器的 VNF 软件,VNF 实现 BRAS 的 DPI、NAT 和 CDN(内容分发网络)等功能板卡,通过 BRAS 的用户流量会被引导进入各个 VNF 组成的 NFV 业务链,这种模式有效提升了 IP 网络边缘业务的部署弹性和可扩展性。

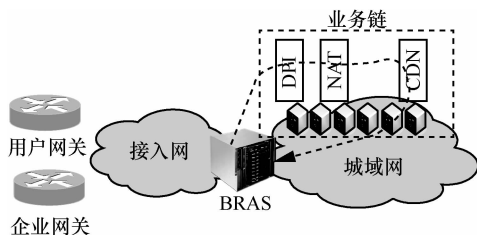


图 1 NFV 业务链在 IP 城域网的应用

NFV 业务链存在串行和星型两种组网方式^[2]：串行连接如图 2 所示，这种网卡端口首尾相连的连接方式结构简单,服务器之间可以不经过交换机直接互连,节约交换机端口,但流量只能从第一台服务器进入,从最后一台服务器流出,因而服务器间的链路带宽可能成为限制网络容量的瓶颈;而在星型连接中,则要求每台服务器保留一进一出两个网卡端口,并上联到 2 个交换机端口,服务器组成业务链时会耗费大量的硬件端口,成本较高。

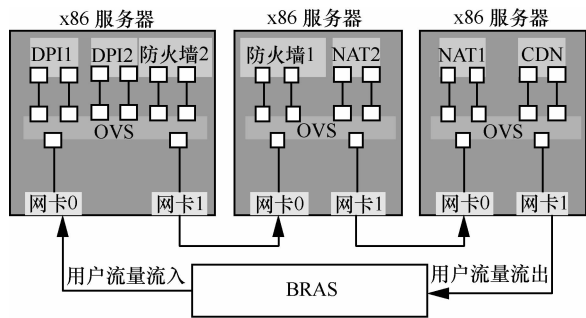


图 2 串行连接方式的业务链

图 2 中,每个 x86 物理服务器均安装开源交换机(OVS),OVS 连接了多个 VM,每个 VM 在实例化以后,形成特定的 VNF,不同的 VNF 通过 OVS 构建的虚拟化网络连接,形成不同的 VNF 业务链。业务链中的每个 VM 只能安装一个 VNF,多个 VM 安装的 VNF 可以相同,并且每台服务器上能够开启的 VM 个数受限于服务器的硬件配置。

本文主要针对串行组网业务链的资源调度问题进行研究,主要原因有 3 个方面:目前在实际工程应用中,以小规模 NFV 业务链居多,成熟的实现方法主要采用串行方法^[3];串行组网可以明显节省硬件成本;在星型组网情况下,流量从 BRAS 设备到达任何 VNF 均为一跳距离,资源分配可采用常规的负载均衡方法解决,而串行方式需要防止流量在前后两台服务器之间进行折返,资源调度问题更为复杂。

2.2 NFV 业务链资源分配问题研究现状

NFV 业务链资源调度的目标是为特定的 VNF 选择一个 VM(VNF 放置),并决定如何将用户流量引导到该 VM(选择业务路径),既要考虑这些流量到达 VNF(VNF)将占用的网络带宽,也要考虑 VNF 处理这些流量需要消耗的计算资源,这是密不可分的两个过程。然而,业界缺乏针对 NFV 场景的业务链资源调度方案,更多是以数据中心业务为场景的单一 VM 放置问题研究,或者单一的 VNF 业务路径选择问题研究(不考虑 VM 资源),下面进行简要分析。

2.2.1 VM 放置问题

VM 放置问题是指在计算资源池中确定多少个 VM 能够安装相同的某类业务,如 Web 服务器,并确定在哪台服务器的具体 VM 上安装上述业务的问题。VM 的放置与 NFV 中的 VNF 放置有着本质的区别,前者用于云 IDC 中,VM 内部可以没有用户流量,因此不消耗带宽资源,而后者用于广域网中,VNF 必须用于处理各类用户流量,从而消耗带宽、CPU 和内存 3 类异构资源资源。

Fumio Machida^[4]希望使用最少的物理服务器实现 VM 冗余放置的问题,利用排队论模型计算出每个业务功能的平均响应时间、需要的 VM 个数和服务器个数,并提出了一种可以保证相同业务放置在不同服务器上的 VM 放置方法,从而实现高可靠性保障,但该方法没有考虑对链路带宽的优化使用。

参考文献[5]提出了一个基于数据中心的放置方案,它考虑了 VM 资源(CPU、内存)、可用性(相关或非相关)和通信资源(带宽、时延等)3种限制,可以对假定的通信资源和 VM 资源进行优化。但是,该方案仅仅考虑了两个 VM 之间的带宽和容量,无法用于端到端多跳的业务链资源计算。此外,方案中引入了多层判断机制,并混合使用了多种算法,因此在工程实施中具有一定复杂性。

2.2.2 业务链选路问题

业务链选路问题是指如何引导用户流量按照正确的顺序穿过指定 VM 的问题,引导的依据可以是距离、时延或者某类资源的最小化要求。

参考文献[6]提出了一个在传统电信网络中的业务链 MiddleBox(专用业务功能设备)放置方法与业务链选路方法,方案根据所有用户流量的业务链序列,构造一个 VNF 依赖关系图,图的顶点是 MiddleBox,边是业务链序列中相邻 MiddleBox 间的连线。将度数最大(即连线最多)的 MiddleBox 放置在离用户入口最近处,并通过求取 MiddleBox 对之间的最短路径之和获取业务路径。该方案适用于基于传统网络设备构造的业务链场景,没有涉及 NFV 场景下计算资源(CPU、内存等)的消耗;此外,在实际应用中,获得所有用户流量关于业务链序列的先验信息比较困难,无法在实际工程中应用。

Hyeonseok Oh^[7]提出的路径选择方法总共分为两步:首先,确定业务链前两个 VNF 的 VM,假设 VNF1、VNF2 均对应多个可选 VM,那么应选择相距最近的 2 个 VM 作为业务链路径必经点;接着,遍历剩下 VNF 之间的可能路径,选择其中的最短路,以上两个步骤的路径加起来便是完整的业务链路径。上述方案仅考虑了对路径长度的优化,没有兼顾 VM 内存、CPU、链路带宽等网络资源。

2.2.3 现状分析

NFV 业务链资源调度需要综合考虑 VNF 放置和 VNF 选路两类问题。目前,业界缺乏针对 NFV 资源(计算与网络)的业务链优化选路技术,存在的问题主要包括以下方面。

首先,单纯云计算资源池中的 VM 放置优化方法应用

于 NFV 业务链时,由于缺乏网络资源控制,会造成 VNF 链路局部拥塞、系统容量低、接入用户少等系列问题;而纯粹基于距离、时延和带宽的业务链选路方法,由于不考虑 VNF 对计算资源的消耗,会造成部分服务器迅速过载,最终导致业务链系统整体性能劣化。

其次,上述各类研究基本只考虑了 VNF 的静态初始化放置,不涉及运行期间 VNF 的动态添加。业务链选路应该建立在 VNF 初始化放置的基础上,并随着系统负载的实时变化进行动态调整。

最后,上述各类方案均较复杂,难以满足工程应用中的业务快速自动部署需求。

目前工程实践中唯一实用化的解决方案,在 OpenDaylight^[3]开源社区的 SFC(service function chain)项目中被提出,它提供了两种 NFV 业务链资源分配方案:最小负荷(least load,LL)和随机(random)分配,这两种方案将 VNF 放置和路径选择结合在一起考虑,简单易行,但是方案同样会造成串行路径折返和链路带宽浪费,同时也没有考虑 VNF 的动态调整问题。因此,对比和完善 OpenDaylight 社区方法,研究实用化的新型资源分配方案,具有重要的工程价值。

综上所述,本文提出了两阶段优化(two-level optimization,TLO)方案,通过 VNF 初始化放置、资源检查和路径选择 3 个关键步骤实现业务链资源的优化分配,同时,方案考虑了 VNF 在服务器上的初始化放置和运行时期 VNF 的动态增加,可以发现全局最优路径,并防止业务链路径折返,大幅提升 NFV 业务链的整体吞吐量。

本文的后续章节介绍了提出的串行 NFV 业务链资源调度方案和 TLO 算法,随后,与 OpenDaylight 提供的标准 LL 和 random 算法进行了实验对比分析。相关算法已经在 OpenDaylight 控制器中进行了集成和验证,可以从 OpenDaylight 社区下载相关代码。

3 串行业务链资源分配方案研究

3.1 方案基本框架

假定整个 NFV 串行业务链是一个只进不出的封闭系统,每到达一批用户将消耗一部分业务链资源,包括 CPU、内存以及业务路径所经链路的带宽。用户流量消耗的资源将被永久扣除,直至资源耗尽无法再接待用户。TLO 方案的目标为吞吐量最大,即资源耗尽时接待的用户数最多。为达到该目标,考虑对物理服务器间的链路带宽、VM 的内



存资源和 CPU 资源这 3 种网络资源进行优化利用。

方案总体分为两个阶段:初始化阶段和运行阶段。初始化阶段是在服务器上预先放置 VNF,初始化阶段完成后便可以对用户进行服务,每个用户都有各自的业务链需求,如用户 A 的需求是 DPI→FW→NAT,用户 B 的需求是 DPI→CDN;运行阶段需要为用户选择合适的业务链路径,使得网络资源被合理利用。图 3 展示了方案基本思路。

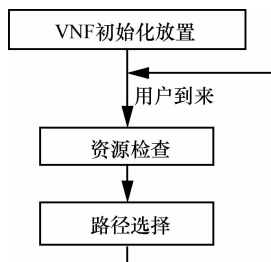


图 3 方案基本思路

3.2 方案描述

方案由如下几个关键步骤组成。

(1)VNF 初始化放置:在服务器上预置各个 VNF。通过经验值或历史数据获得每个 VNF 的用户需求百分比,按占比数由大到小的顺序依次在不同服务器上放置一个 VNF。假设需求百分比排序结果为 DPI>FW>NAT,那么应该在服务器 1 上开启 DPI,在服务器 2 上开启 FW,在服务器 3 上开启 NAT。本方案使 VNF 到用户入口的距离反比于 VNF 热度,避免用户流量在初始拓扑中出现路径折返,浪费链路带宽。

(2)资源检查:每当用户流量到来时,检查业务链上 VNF 对应的网络资源是否足够。若资源不足,通过新增 VM 创建新的 VNF 资源,实现 VNF 的动态增加。新增 VM 的方法和具体流程如图 4 所示。

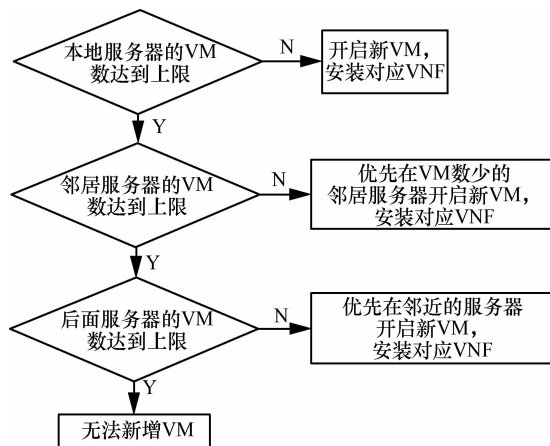


图 4 VNF 新增具体流程

(3)路径选择:本文提出了一套业务链路径选择算法,具体实现流程如图 5 所示。

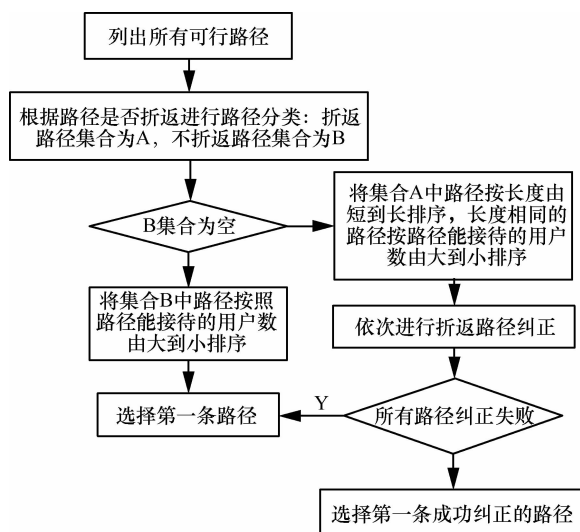


图 5 路径选择算法具体流程

已知用户对 VM 资源的消耗量,业务链路径能接待的用户数计算方法如下:

- 路径接待用户数= $\min(\text{VNF1 接待用户数}, \dots, \text{VNFn 接待用户数})$;
- VNF 接待用户数= $\min(\text{CPU 接待数}, \text{内存接待数})$;
- CPU 接待数=VM 的 CPU 剩余资源/用户对 CPU 的消耗量;
- 内存接待数=VM 的内存剩余资源/用户对内存的消耗量。

另外,算法还提供了用于服务器间链路带宽优化的折返路径优化机制。

折返路径的判断:串行连接方式下,不折返路径长度为恒定值(服务器个数+1),因此大于此数值的路径必然折返。

折返路径的纠正:首先,将可行路径按照长度由短到长、接待用户数由多到少的顺序进行排序;然后,从第一条路径开始,通过新增 VM 的方式纠正折返路径。VM 新增方法要遵循两个原则:第一,新增 VM 个数最少;第二,在可以放置新增 VM 的服务器中,选择当前 VM 个数最少的服务器。一旦成功纠正,立即停止纠正机制,并选择纠正成功的路径作为业务链路径;若所有路径都纠正失败,那么选择长度最短的路径作为业务链路径。

算法的路径折返纠正机制实现了对物理服务器间的链路带宽的优化利用,并且在进行路径选择时,充分考虑了对 VM 内存和 CPU 资源的均衡使用。

4 实验与分析

将 TLO 方案与 OpenDaylight 提出的 LL 和 random 方案进行比较。

LL 和 random 方案均采用随机的 VNF 初始化放置方法。LL 方案在进行路径选择时,每次均选择能够接待用户数最多的 VNF(负荷最小的 VM),然后将这些 VNF 虚拟机串接起来构成最终的业务路径;而 random 方案则随机选择 VNF,将这些 VNF 虚拟机串起来构成最终路径。

进行两组实验,分别对 TLO 方案的两个阶段方法进行性能分析。实验 1 将 TLO 初始化方法同随机初始化法进行对比分析;实验 2 在采用 TLO 初始化方法的前提下,将 TLO 路径选择算法同 LL 及 random 路径选择算法进行比较。

实验输入为随机生成的业务链序列,每个业务链序列描述了一个用户流量需要通过的 VNF 以及通过顺序,业务链长度假定为 3~5 个 VNF。实验将依次为用户流量找到合适的业务链路径,每服务完一位用户才能接待下一位用户。

实验设置了 10 台物理服务器。根据物理服务器实际测试结果,OVS 会消耗近 50%的服务器资源,因此假设每个服务器最多能开启 4 台 VM 是较为合理的。另外,实验假定有 7 个 VNF,VNF 的相关参数见表 1。

由于 IP 网络边缘的物理服务器间链路带宽和 VM 资源比例关系大致为 10:1,因而假设物理服务器间链路带宽初始值为 10,VM 的 CPU 和内存资源初始值为 1。

两组实验各自重复 400 次,每次使用不同输入,最后对网络平均吞吐量和 VNF 间平均路径长度进行统计,这两个性能指标定义如下:吞吐量=接待用户数;VNF 间路径平均长度=路径长度/业务链长度。

4.1 实验 1

随机初始化方法将 7 个 VNF 随机排序,并逐一放置在各个服务器上。由于不同排序结果会导致不同的实验结果,因此选取 4 种随机排序同 TLO 初始化方法进行比较。由于初始化方法只是对 VNF 的预置,需要运行以后才能比较方法性能,因此运行阶段的路径选择方法采用 TLO 路径选择算法,实验结果如图 6、图 7 所示,其中图 6 纵坐

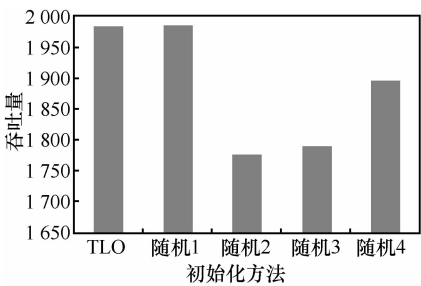


图 6 不同初始化方法吞吐量对比结果

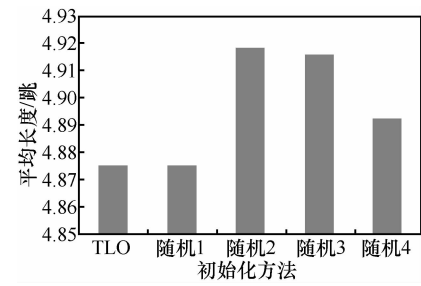


图 7 不同初始化方法 VNF 间路径平均长度对比结果

标的单位为用户个数/session。

4.2 实验 2

实验 2 的结果如图 8 和图 9 所示,其中图 8 纵坐标的单位为用户个数/session。可以看出,虽然随机初始化 1 的

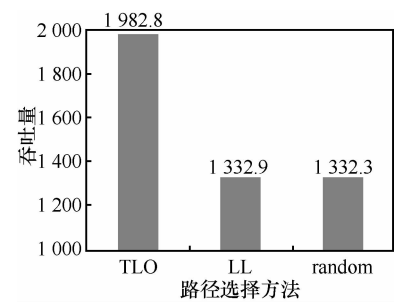


图 8 不同路径选择算法的吞吐量对比

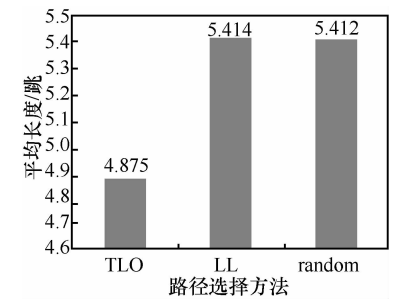


图 9 不同路径选择算法的平均路径长度对比

表 1 实验所用的参数

参数名	VNF1	VNF2	VNF3	VNF4	VNF5	VNF6	VNF7
用户对 VNF 的需求占比	40%	30%	10%	8%	5%	5%	2%
用户对 VM CPU 消耗量	0.2%	0.3%	0.2%	0.1%	0.4%	0.3%	0.1%
用户对 VM 内存消耗量	0.04%	0.06%	0.03%	0.01%	0.1%	0.07%	0.03%



性能与 TLO 初始化方法不相上下,但随机初始化方法的平均性能明显不及 TLO 方法。另外,TLO 路径选择算法的总体性能优于 LL 和 random 算法。在本实验中,TLO 业务链资源分配方案提升了 48.7% 的网络吞吐量,降低了 11.1% 的 VNF 间路径平均长度。

5 结束语

本文首先详细分析了 NFV 业务链资源分配问题研究现状,然后提出了 TLO 业务链资源分配方案,该方案分为初始化和运行两个阶段。在初始化阶段,TLO 方案提出了一种简单易行的 VNF 初始化放置方法;在运行阶段,TLO 方案考虑了对 3 种网络资源(物理服务器间链路带宽、VM 的 CPU 资源以及 VM 内存资源)进行优化利用,帮助用户流量选择最合适的业务链路径。最后通过实验,验证了 TLO 方案在上述两个阶段的性能优化能力。

本方案支持 IP 网络边缘网络功能的分布式部署,实现了 NFV 基础资源的集约化利用,有效增大了 NFV 设施的承载用户数。同时,网络功能在缩容、扩容方面更富于弹性,可以依据网络业务和用户流量情况进行按需配置,增加了业务链部署和管理的灵活性。

随着业界对 NFV 业务链研究的不断深入,运营商关注的重点也从功能场景逐步过渡到业务链的各类实现技术,包括业务链组网、业务链资源分配、业务链自动部署与管理技术等方面,本文的研究将有助于 NFV 业务链技术推广部署。本文的资源分配方案尚属于初步阶段,后续工作包括修正用户流量模型、完善资源感知、批量优化用户资源等几个方面。

参考文献

- 1 John W, Pentikousis K, Agapiou G, *et al.* Research directions in network service chaining. Proceedings of 2013 IEEE SDN for Future Networks and Services (SDN4FNS), Trento, Italy, 2013: 1~7
- 2 唐宏, 欧亮. 网络功能虚拟化中的网络转发性能优化技术研究. 电信科学, 2014, 30(11): 135~139

Tang H, Ou L. Research on forwarding performance for network function virtualization. Telecommunications Science, 2014, 30(11): 135~139

- 3 OpenDaylight SFC Project. https://wiki.opendaylight.org/view/Service_Function_Chaining:Main, 2014
- 4 Machida F, Kawato M, Maeno Y. Redundant virtual machine placement for fault-tolerant consolidated server clusters. Proceedings of 2010 IEEE Network Operations and Management Symposium (NOMS), KIX, Japan, 2010: 32~39
- 5 Jayasinghe D, Pu C, Eilam T, *et al.* Improving performance and availability of services hosted on iaas clouds with structural constraint-aware virtual machine placement. Proceedings of 2011 IEEE International Conference on Services Computing (SCC), Washington DC, USA, 2011: 72~79
- 6 Zhang Y, Beheshti N, Beliveau L, *et al.* StEERING: a software-defined networking for inline service chaining. Proceedings of IEEE International Conference on Network Protocols, Göttingen, Germany, 2013: 1~10
- 7 Oh H, Yu D, Kim N, *et al.* Study on search algorithm for finding an efficient service-chaining path. Advanced Science and Technology Letters, 2014(51): 102~105

[作者简介]



唐宏,男,中国电信股份有限公司广州研究院数据通信研究所所长,主要从事 IP 承载网、下一代互联网、网络新技术方面的研发与管理工作。



罗雨佳,女,中国电信股份有限公司广州研究院数据通信研究所工程师,主要从事 IP 承载网络技术、SDN/NFV 技术相关研究等工作。