



基于 SDN 的按需智能路由系统研究与验证^{*}

王歆平¹, 王 茜², 刘恩慧¹, 马亦然³, 林程勇¹, 史 凡³

(1. 华为技术有限公司北京研究所 北京 100095; 2. 中国电信集团公司 北京 100032;
3. 中国电信股份有限公司北京研究院 北京 100035)

摘 要: 软件定义网络(SDN)是一种新型的网络技术, SDN 开放可编程的理念与网络管道智能化的需求能够完美地契合。在基于 SDN 的下一代高智能网络(S-NICE)总体架构研究的基础上进行探索, 论述了 S-NICE 架构中的核心系统——按需智能路由系统的关键技术、系统设计、部署验证和应用场景。

关键词: 软件定义网络; 智能增强型网络; 基于 SDN 的下一代高智能网络; 按需; 智能路由

doi: 10.3969/j.issn.1000-0801.2014.04.002

Research and Verification on SDN-Based On-Demand Smart Routing System

Wang Xinping¹, Wang Qian², Liu Enhui¹, Ma Yiran³, Lin Chengyong¹, Shi Fan³

(1. Beijing Research Center, Huawei Technology Co., Ltd., Beijing 100095, China;
2. China Telecom Corporation, Beijing 100032, China;
3. Beijing Research Institute of China Telecom Co., Ltd., Beijing 100035, China)

Abstract: SDN is a new network technology. Open programmable SDN concept can be perfectly combined with intelligent network pipeline demand. Research on S-NICE architecture has explored in this field. The on demand custom routing system which is a kernel system in the S-NICE architecture was described from several aspects, such as key technologies and system design, deployment and verification, as well as scenarios.

Key words: software defined networking, network intelligence capability enhancement, S-NICE, on-demand, smart routing

1 引言

智能型通信网络^[1]是指客户感知良好、运营管理方便、业务开通灵活的智能增强型网络, 提供用户高速协同接入、需求感知分析、资源自助指配、质量按需保障的差异化服务。随着信息技术的快速发展和用户需求的不断变化, 网络正在由传统粗放型“哑管道”向精细化“智能管道”转型。基于控制转发分离和开放可编程理念的 SDN^[2]新型网络技术不断成熟, 用于满足智能管道“按需发放”、“流量优

化”、“感知”、“开放”、“协同”的需求^[3], 比传统方法更具优势。SDN 控制器集中拥有全网视角的网络拓扑, SDN 流控算法按需计算流路径、管理调度流量, 可以满足流量优化的需求; SDN 控制器的集中拓扑和开放北向接口, 便于用户按需实时定制业务, 可以满足按需发放的需求; 通过 OpenFlow (OF) 等控制通道协议, 转发面将网络状态、流量实时上报到控制器进行流量分析, 可以满足感知的需求; SDN 控制器开放北向可编程接口, 便于第三方应用的快速生成, 可以满足开放的需求; SDN 集中控制面之后, 便于不

^{*} 国家发展和改革委员会“可演进的下一代高智能网络架构研究和试验”课题

同区域网络间的协同,可以满足协同的需求。

本文在基于 SDN 的下一代高智能网络(S-NICE)的总体架构研究的基础上,重点分析设计了基于 SDN 技术的按需智能路由系统,通过关键技术的研究和原型系统的实现及实验验证,不仅探索了如何实时、自助、按需定制地为 VIP 用户或应用开展智能路由服务,而且也探索了如何平滑地从传统 IP 网络向 SDN 过渡,为将 SDN 技术应用到运营商智能管道网络中做出了有价值的研究和范例。

2 基于 SDN 的下一代高智能网络方案

2.1 网络架构

基于 SDN 的下一代高智能网络(如图 1 所示)分为承载控制层和承载层。承载控制层向上层应用提供满足业务需求并与承载层能力相匹配的管理控制能力,提供基于感知分析的 SDN 控制和策略控制。承载层满足各类用户和终端的灵活接入、各类应用的按需路由优化和策略执行,并具备应用可感知、用户可区分、路由可优化的新型转发体系。

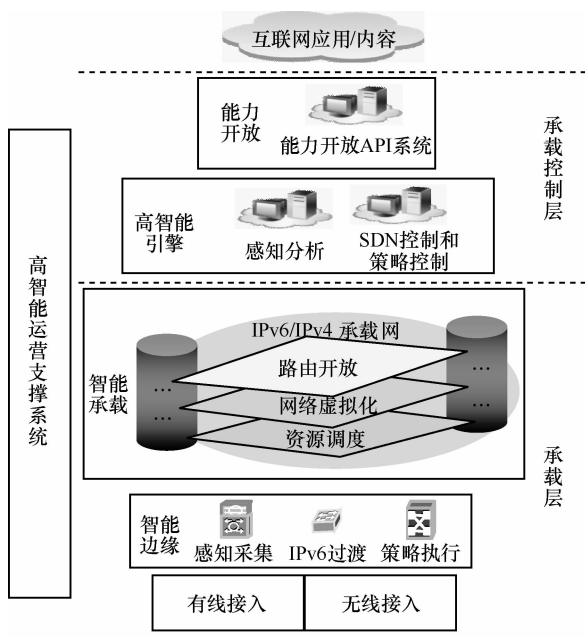


图 1 下一代高智能网络架构模型

从实现的角度,下一代高智能网络架构由 5 个系统组成:能力开放系统、高智能引擎系统、软件定义的智能承载系统、智能边缘接入系统、高智能运营支撑系统。

- 能力开放系统: 为上层应用提供统一的开放接口,支持智能型网络中业务和网络能力的调用,也支持对智能承载、智能边缘接入系统中资源、路径、服务等能力的灵活调用和管理。

- 高智能引擎系统: 面向感知的高智能引擎系统具体包括感知分析系统、SDN 控制和策略控制等子系统。通过对承载层各层面信息的统计分析,针对不同的用户和应用需求,设置相应的资源调用策略或者应用路由。通过 SDN 控制和策略控制功能,选择适合不同应用流量模型的路由、存储或处理资源,保障业务承载的质量并提升承载网络的传送效率。
- 软件定义的智能承载系统: 是新型的基于 SDN 路由开放技术、网络虚拟化的新型路由系统,支持对承载网的传送路由资源的虚拟抽象信息提供,支持与高智能引擎系统协同的可编程、可扩展的应用转发能力。
- 智能边缘接入系统: 支持用户和应用接入信息的感知采集、海量终端的无缝接入、IPv6 过渡技术等,为构建端到端的下一代高智能网络提供用户可识别、应用可区分、跨网协同的智能边缘控制能力。
- 高智能运营支撑系统: 支持与 SDN 控制和策略控制相关的针对用户和应用的认证鉴权、计费账务、审计管理、业务运营管理等功能。

2.2 按需智能路由系统特征

基于下一代高智能网络架构模型,按需智能路由系统由 SDN 控制器和智能承载设备构成(如图 1 所示)。按需智能路由系统用 SDN 技术为 VIP 用户或应用按需定制、自动实时提供有质量保证的流转发路径,既可以实现管道的按需货币化,也可以提升 VIP 用户的应用体验,满足精细化运营管道的智能管道需求。

按需智能路由系统特征如图 2 所示,在按需智能路由服务中,80%的普通用户报文通过传统 IP 转发,而 20%的 VIP 用户,根据自己的需要,付费定制高质量的服务,享用优质 QoS 保证的 OpenFlow(OF)流转发路径;SDN 控制器上的流控算法按需为 VIP 付费用户计算定制路径,可以支持粗或细粒度、静态或动态 VIP 流量识别和 NaaS 定制路由服务;当 SDN 控制器出现故障时,不影响传统普通业务的 IP 转发,已经定制了按需定制路由服务的 VIP 用户,依旧享有转发设备上已经建立好的精细化流转发路径,而新的 VIP 用户也通过传统 IP 转发,不影响现网应用。

3 按需智能路由系统设计和实现

如图 3 所示,按需智能路由服务是一个端到端的业务方案,为了实现该业务功能,对 SDN 控制器和转发设备提出了许多新的能力需求。SDN 控制器需要支持开放可编程

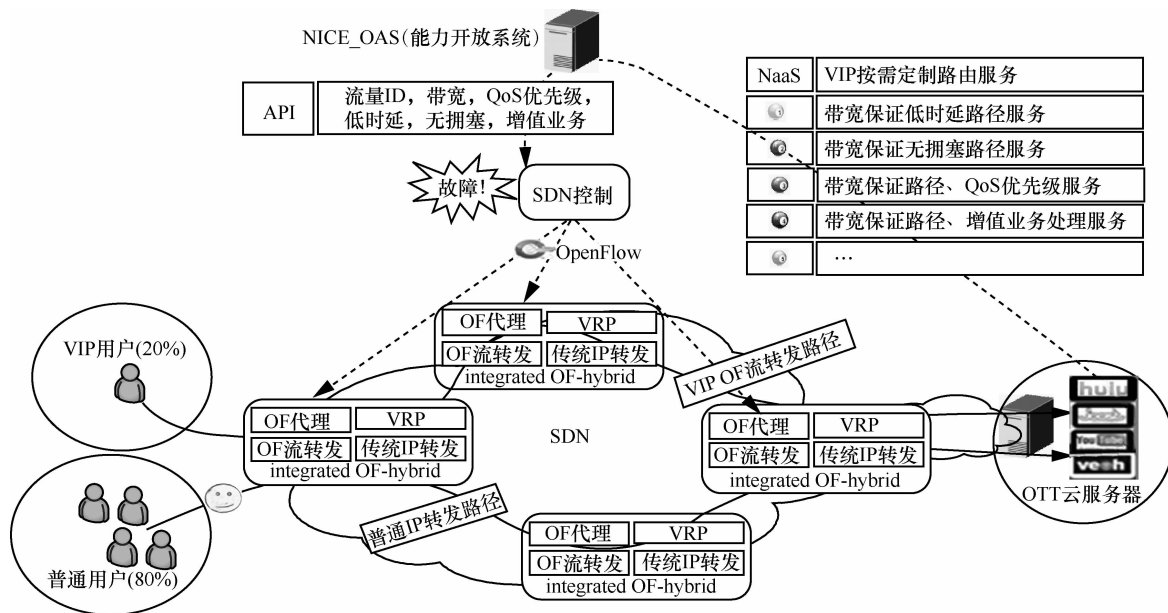


图 2 按需智能路由系统特征

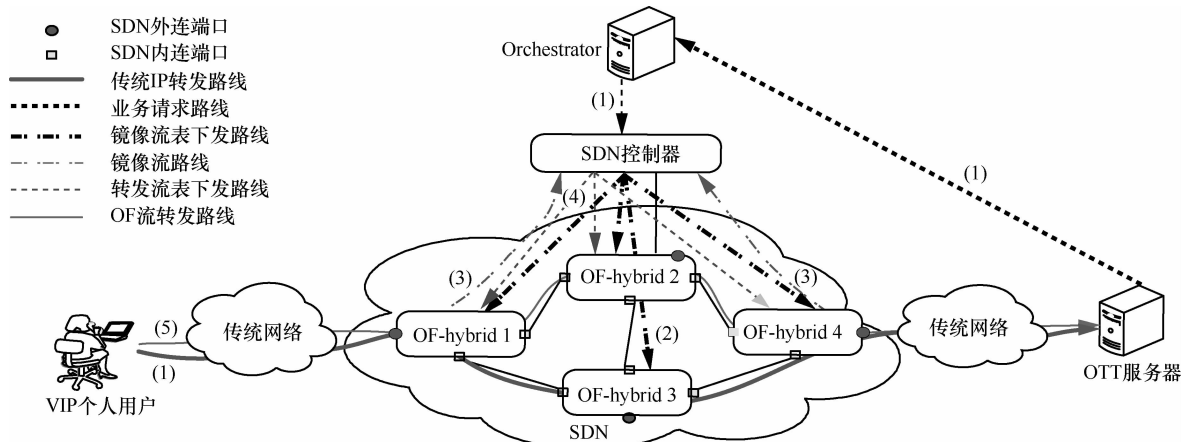


图3 按需智能路由系统架构和处理流程

的北向接口 (northbound interface, NBI)、流路径计算、拓扑收集、SDN 域边界节点识别、按需镜像流机制和 4~7 层按需镜像流机制等关键技术,而且转发设备需要既支持传统的 IP 转发,又支持 OF 流转发。

3.1 系统设计

(1)SDN 控制器

SDN 控制器平台提供 OF 的各项基础通信和管理功能,目前业内有很多开源的 SDN 控制器平台,如 Floodlight^[4]、POX^[5]、NOX^[6]、ODL^[7]等,本原型使用了一款华为控制器平台,基于通用服务器平台、C 语言和 Linux OS 开发。如图 4 所示,NICE 组件是实现按需智能路由服务各项特定功能的关键组件,它作为一个应用组件可以集成在任一款 SDN 控制器平台上。

图 4 展示了 NICE 组件接口和内部功能模块。组件接口包括北向接口(传递定制参数 API)、OF-packet-in/out 接口(用于传递 packet-in 或 packet-out 消息)、流表接口(用于下发镜像流表和转发流表)、拓扑接口(用于获取 SDN 域网络拓扑)以及与增值业务组件之间的接口(用于和增值业务组件通信,按照定制需求,实现增值业务功能)。NICE 组件的功能模块主要包括边界识别、策略管理、应用感知、流管理 & 流表计算、流路径计算模块。边界识别模块根据 packet-in 消息中携带的 in/out 接口信息和接口属性信息(内外联、上下行信息),确定 SDN 边界节点,并保存到流管理表中;策略管理模块管理从北向接口接收到的策略参数,完成策略组合识别、策略仲裁,并保存流策略到流管理表中;流管理 & 流表计算模块管理、维护流管理表,向流路

径计算模块请求路径计算,根据路径和策略计算流表、流表下发;流路径计算模块根据拓扑,使用SDN路径计算算法,计算有质量保证的流转发路径;应用类型感知模块对上送到控制器的流,进行深度感知分析,实现针对符合要求的特定应用,进行流转发流路径的计算。

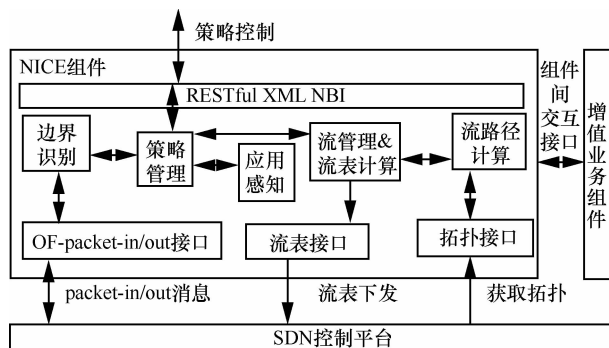


图4 NICE 组件架构

(2)北向接口设计

北向接口基于标准的RESTful接口和规范,根据按需智能路由业务的需求,定义了可编程的API参数和XML消息类型格式。

北向接口包括两类参数:标识参数和策略参数。其中标识参数可以包括流量标识和应用感知类型标识,流量标识用来标识一条流,可以是1~5元组(源IP地址、目的IP地址、协议号、源端口号、目的端口号);应用类型标识用于标识承载在一条流里的一个4~7层的应用类型,例如HD video、P2P、VoIP等。而策略参数是提供给SDN流路径算法按需计算流转发路径用的,可以包括带宽(如2 Mbit/s、4 Mbit/s、8 Mbit/s、10 Mbit/s)、QoS等级(如金、银、铜)、低时延等级(如普通、高级)、无拥塞等级(如普通、高级)、增值业务标识(如URL过滤、流量分析)。

XML消息格式被设计用于发起业务控制、封装和参数传递。设计了4类基本的消息接口:业务创建、删除、变更消息接口和SDN控制器能力上报消息接口。其中,业务创建消息接口用于实时创建VIP按需定制服务,包括业务创建消息和创建应答消息;业务删除消息接口用于实时删除VIP按需定制服务,包括业务删除消息和删除应答消息;业务变更消息接口用于实时变更VIP按需定制服务,包括业务变更消息和变更应答消息;如SDN控制器能力上报接口用于将SDN的能力上报给编排设备Orchestrator,包括上报消息和上报应答消息。此外,根据需要,也可以定义其他的附加消息接口。

(3)OF-hybrid 混合转发设备

转发设备采用OF-hybrid路由器,基于华为NE40E平台开发,可以同时支持传统路由转发和OF流转发。在转发面,通过配置使能OF流水线(pipeline)转发功能。当未使能该功能时,路由器只是一台普通的传统路由器。OF-hybrid混合转发设备的端口属性分为两类:SDN外连端口(SDN-ExtPort)和SDN内连端口(SDN-IntPort)。如图3所示,SDN外连端口用小圆圈表示,代表通过该端口可以与传统路由器相连;SDN内连端口用小方块表示,代表通过该端口可以实现SDN域内各转发设备之间的互连。接口属性可以通过配置预先指定。

当数据流到达OF-hybrid设备时,将首先触发流表查找,如果未匹配,则会有3种miss-table动作处理:上送SDN控制器、IP转发或者丢弃。需要预先在SDN外连端口和SDN内连端口上配置miss-table动作为IP转发。

3.2 关键技术

本节重点介绍实现按需智能路由服务所需的关键技术,SDN/OpenFlow规范的基础通用功能此处不再赘述。

(1)integrated OF-hybrid 混合转发技术

integrated OF-hybrid混合转发技术是OF-hybrid路由器的核心机制,通过配置使能OpenFlow流水线功能,从而激活路由器,同时既支持传统路由转发,又支持OpenFlow流转发。如图5所示,当使能OpenFlow转发时,报文进入路由器后,先进行OpenFlow流表的匹配查找,如果匹配,则先进行OpenFlow的流表处理,再进入传统流水线查表转发,如果未匹配,则进行table-miss动作:上送控制器、继续IP转发或者丢弃处理。但是如果未使能OpenFlow转发,报文进入路由器后,直接进入传统流水线完成转发。

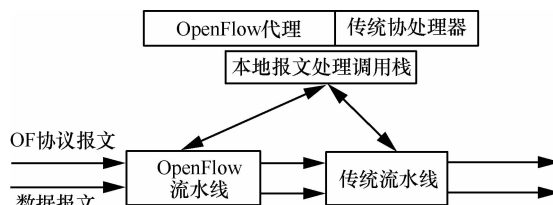


图5 integrated OF-hybrid 混合转发技术

(2)流转发路径计算

SDN流路径计算功能是流量调度的核心,根据用户的定制需求,基于SDN网络拓扑,计算一条从SDN域入口到出口节点的流路径。为了实现这个功能,要求满足以下条件:流路径计算算法、SDN网络拓扑、SDN域入出边界节点



和用户流量调度策略参数。流路径算法可以使用 CSPF (constrained shortest path first) 来计算一条符合约束条件的流路径。用户流量调度策略也就是用户定制参数,能够通过北向接口获取。

因为 integrated OF-hybrid 转发设备同时支持 OpenFlow 协议和传统的 IGP (如 OSPF), 可以有两种方法收集 SDN 域网络拓扑。一种是通过 ONS 规范的改进 LLDP (link layer discovery protocol) 实现拓扑收集;另一种方法是充分利用传统 OSPF 收集拓扑,然后镜像到 SDN 控制器上。

(3) 按需镜像流表机制

为了识别 SDN 域边界节点,需要发送用户流量到 SDN 控制器,但是这将给 OpenFlow 控制通道造成很大的负载压力,给控制器造成很大的性能处理压力。为了解决这个问题,设计了按需镜像流表机制。镜像流表是指查找流表匹配后,动作是 mirror 的流表。动作 mirror 意味着这条匹配流不但继续进行传统的 IP 转发,同时将被复制一份发送到 SDN 控制器。

从北向接口收到标识参数后,SDN 控制器将发送镜像流表到所有的 SDN 边界节点。如果一条数据流匹配了镜像流表,它将被发送到 SDN 控制器。

此外,如果镜像流表长期不被使用,能够通过老化机制,定期自动删除。

(4) SDN 边界节点识别技术

当一条用户数据流穿过 SDN 时,它首先到达 SDN 的边界入口节点的 SDN 外连端口,触发查找流表,此时如果存在镜像流表,OF-Hybrid 设备除了继续 IP 转发该条数据流,还要复制一份,用 OpenFlow packet-in 消息封装后发送到控制器。SDN 控制器收到后,根据 packet-in 消息中携带的端口号和上行属性信息,断定该节点是这条流在 SDN 域的边界入口节点。

此外,该数据流通过 SDN 域的内部中间节点继续 IP 转发,因为 SDN 内连端口被配置为 miss-table 动作作为 IP 转发,而且此时中间节点还没有 OpenFlow 流表,所以查表后进行 table-miss 动作,继续 IP 转发。

最后,该数据流到达 SDN 域边界出口节点,SDN 外连端口触发查找流表,此时如果存在镜像流表,OF-hybrid 设备除了继续 IP 转发该条数据流,还要复制一份,用 OpenFlow packet-in 消息封装后发送到控制器。SDN 控制器收到后,根据 packet-in 消息中携带的端口号和下行属性信息,断定该节点是这条流在 SDN 域的边界出口节点。

至此,得到了该条流在 SDN 域的边界入口节点和出口节点。

(5) 应用类型识别技术

为了给承载在一条流中的某一个 4~7 层应用提供高质量的流转发路径服务,需要在 SDN 控制器上进行应用感知识别,这个功能可以用现有的 DPI (deep packet identification) 等技术来实现。应用类型识别功能是一个高级功能,通过北向接口发来的应用感知类型标识参数来使能,如果该参数为有效值,则使能应用感知功能,如果为无效值,则不使能。

3.3 系统处理流程及优势

根据图 3,典型按需智能路由服务的处理流程设计为以下几个步骤。

(1) VIP 个人用户根据自己的需求,通过传统 IP 转发路径,提交定制路由服务申请到 OTT (over the top) 的应用服务器。定制参数可以是用户业务流的流量标识参数和策略参数,其中流量标识参数可以是五元组标识、需要感知的 4~7 层应用类型标识等,策略参数可以是带宽、QoS 等级、低时延等级、无拥塞等级、增值业务标识 ID 等。收到用户定制请求后,OTT 服务器向编排设备 Orchestrator 继续发送用户定制路由服务请求,Orchestrator 收到请求后,根据北向接口规范封装定制参数,并发送给 SDN 控制器。

(2) SDN 控制器通过北向 API 收到定制路由服务的流量标识和策略参数后,向所有 SDN 域的边界节点下发该用户业务流的镜像流表。

(3) 用户业务流经过 SDN 域入、出边界 integrated OF-hybrid 转发节点时,查镜像流表,上送到 SDN 控制器,SDN 控制器识别出入口边界节点和出口边界节点。

(4) SDN 控制器根据识别出的入口边界节点和出口边界节点以及北向 API 发来的策略参数信息,使用路径算法,为该 VIP 用户计算符合需求的高质量保证的路径和转发流表,并将转发流表通过 OpenFlow 控制通道分别下发到该路径上相应的各个 integrated OF-hybrid 转发节点上。

(5) VIP 用户通过按需付费,享受定制的、高质量保证的流转发路径服务,提高了使用体验。

更进一步,如果定制业务需要做应用感知,即从北向接口发来的定制参数中,应用感知类型标识参数被设置为有效值,则在业务处理流程的步骤(3)中,收到镜像数据流后,SDN 控制器需要首先进行 4~7 层的应用类型识别,如

果识别出的应用类型与应用感知类型标识的应用类型相同,则SDN控制器继续识别入口边界节点和出口边界节点。

按需智能路由系统具有5个特点:平滑可靠的SDN过渡、按需服务、精细化流转发、流量可调控、节约资源。平滑可靠的SDN过渡是指可以从传统IP网络平滑演进到SDN,这得益于integrated OF-hybrid设备同时支持IP转发和OF流转发,当控制器出现故障时,不影响传统普通业务转发;按需服务是指可以根据VIP用户或VIP业务不同的策略需求,提供高质量、精细化流转发服务;精细化流转发是指可以支持高质量保证的流转发;流量可调控是指SDN控制器自动识别SDN边界节点,SDN流控算法按需计算流转发路径,进行流量控制;节约资源是指按需镜像相比传统首分识别方案,大幅减少控制器与转发设备之间控制通道的压力以及控制器的处理压力。

4 系统验证和应用场景

4.1 实验部署和验证

按需智能路由系统已经在华为实验室(如图6所示)和中国电信江苏省无锡市试验网进行了部署和验证,主要测试和验证内容如下。

(1)通过OpenFlow 1.3协议^[8],基于SDN控制器控制的NICE-SRS系统能够成功建立,控制器和转发设备正常通信。

(2)SDN北向接口通信成功,通过此接口,成功实现S-NICE按需智能路由服务的建立、删除和修改。

(3)SDN控制器能够实时根据用户定制参数,成功计算有高质量保证的流转发路径,提升用户的使用体验。

(4)当SDN控制器发生故障时,传统IP业务不受影响,正常工作;先前申请了S-NICE按需智能路由服务的用户依然享有高质量保证的流转发路径服务。只有后续新的用户不能使用S-NICE按需智能路由服务。

(5)可以同时支持IPv4和IPv6业务的OpenFlow流转发。

4.2 应用场景设计

按需智能路由系统可以根据不同的用户或商业需求,提供许多应用场景。本次实验分别针对VIP个人用户、ICP、企业,设计了3种典型常用的业务场景。

(1)为VIP个人用户提供按需定制路由服务

首先,当VIP个人用户正在观看ICP(internet content provider)提供的高清视频时,若网络质量恶化,则该用户的视频网页会变得有马赛克、不清晰。为了改善观看体验,该用户通过点击填写集成在视频网页上的定制菜单按需申请高质量定制路由服务,定制参数可以是带宽、QoS等级等。不同的带宽和QoS等级对应不同的价格,用户可以根据自己的需要和承受能力,支付相应的费用,享受相应的高质量路径服务。整个操作非常简单、实时。完成定制后,该用户由一个普通用户成为一个VIP个人用户;接着,SDN控制器和网络根据定制策略参数,实时按需为该用户的高清视频流量生成一条带宽保证的高质量流转发路径;最后,该用户访问高清视频网页时,观看过程一直保证清晰流畅,不受网络拥塞影响。

(2)为ICP提供应用感知的定制路由服务

首先,一个ICP希望改善它所有的高清视频用户的观看体验,于是该用户通过登录一个由网络运营商提供的自助服务网页,填写定制参数来按需申请应用感知的定制路

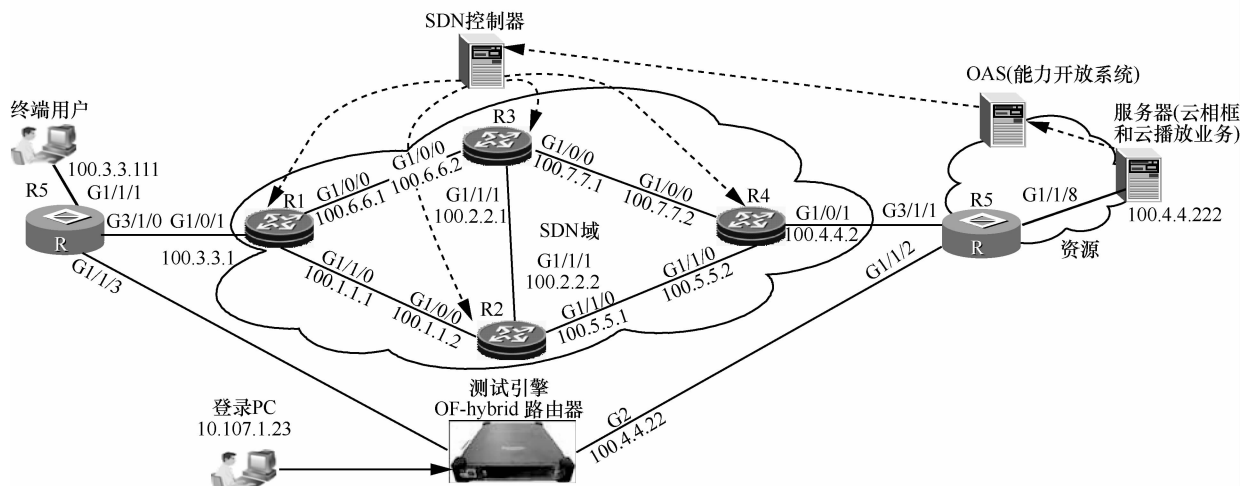


图6 按需智能路由原型系统实验室部署



由服务,对高清视频用户提供高质量的流转发路径服务。完成定制后,这个高清视频应用由一个普通应用成为VIP应用;接着,SDN控制器和网络根据这些定制策略参数,为每个访问该ICP高清视频网站的个人用户,实时生成多条高质量流转发路径;最后,该VIP个人用户访问高清视频网页时,观看过程一直保证清晰流畅,不受网络拥塞影响。

(3)为企业应用感知的高级定制服务

首先,企业通过登录一个由网络运营商提供的自助服务网页,填写定制参数来按需申请“应用感知的高级定制服务”,申请增值业务URL过滤和企业上网流量统计分析;接着,SDN控制器根据定制策略参数,实时下发控制指令,SDN智能边源设备实时执行URL过滤和上网流量分析;最后,企业员工在工作时间访问娱乐网站返回禁止访问提示页面,而且企业管理者可以查看企业上网流量统计分析报表。

5 结束语

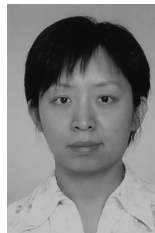
下一代高智能网络架构和按需智能路由系统,既探索了用SDN技术实现智能管道的需求,又探索了从传统IP网络向SDN平滑过渡的过程,通过原型验证其是新颖、有价值、可行的。同时,中国通信标准化协会(CCSA)和国际电信联盟(ITU-T)^[9]都在开展下一代高智能网络架构和系统的标准化工作,以促进相关技术方案的成熟。未来在SDN流量调度算法、SDN控制器的性能和可靠性、转发设备的性能和标准化等方面仍有很多工作需要继续深入。

参考文献

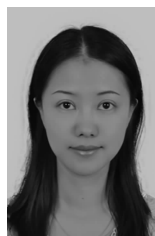
- 1 YD-T 2636-2013.智能型通信网络 总体框架和要求,2013
- 2 ONF. Software-Defined Networking: the New Norm for Networks, 2012
- 3 ITU-T Y.2301. Network Intelligence Capability Enhancement - Requirements and Capabilities, 2013
- 4 BigSwitch networks. <http://tinyurl.com/cgepwdj>
- 5 Meccauley J. POX: a python-based openflow controller. <http://www.noxxrepo.org/pox/about-pox/>
- 6 Gude N, Koponen T, Pettit J, et al. NOX: Towards an Operating System for Networks. ACM SIGCOMM Computer Communication Review, 2008, 38(3):105~110
- 7 Open daylight. <http://www.opendaylight.org/>

- 8 ONF. OpenFlow Switch Specification Version 1.3.1, 2012
- 9 ITU-T Y. S-NICE-reqts. Requirements and Capability Framework for NICE Implementation Making Usage of Software Defined Networking Technologies (S-NICE), 2013

[作者简介]



王歆平,女,华为技术有限公司北京研究所IP技术研究部高级工程师,主要研究方向为软件定义网络/OpenFlow、网络虚拟化、网络开放可编程等。



王茜,女,博士,现就职于中国电信集团公司,多年从事电信宽带网络领域研究工作,在IP网络规划和优化、下一代互联网技术、智能管道技术、网络虚拟化技术等方面有深入研究和积累,并参加ITU-T、IETF、CCSA等标准组织的标准编制工作,发表文章30余篇,出版专著5本。



刘恩慧,女,华为技术有限公司北京研究所IP技术研究部高级项目经理、技术专家,主要研究方向为软件定义网络/OpenFlow、网络虚拟化、网络开放可编程。

马亦然,男,博士,中国电信股份有限公司北京研究院工程师,主要从事光传输与接入、智能管道等方面的研究工作,并参加ITU-T、FSAN、CCSA等标准组织的标准编制工作,在IEEE和OSA等国际期刊以及OFC等国际会议发表论文30余篇。

林程勇,男,华为技术有限公司北京研究所IP技术研究部高级工程师,主要研究方向为软件定义网络/OpenFlow、网络虚拟化、网络开放可编程等。

史凡,男,中国电信股份有限公司北京研究院网络架构与前沿技术研究室主任,MEF(城域以太网论坛)中国工作组联合主席,主要研究方向为下一代互联网、SDN/NFV等。

(收稿日期:2014-03-25)