



专题:下一代互联网

基于 SDN 架构的 IPv6 过渡技术设计与实现

孙 琼¹, 解冲锋¹, 赵慧玲¹, 卢绪山², 刘树成²

(1. 中国电信股份有限公司北京研究院 北京 100035; 2. 华为技术有限公司 深圳 518129)

摘 要:重点研究了基于 SDN 的 IPv6 过渡技术,首次创新性地提出了在运营商网络中引入基于 SDN 的 IPv6 过渡技术的主要应用场景及实施优势、系统实现架构、关键技术及相关部署要求,该方案的提出对于运营商网络中基于 SDN 的 IPv6 过渡技术的引入具有重要的指导意义。

关键词:软件定义网络;IPv6;过渡技术

doi: 10.3969/j.issn.1000-0801.2014.04.003

Design and Realization of SDN-Based IPv6 Transition Technology

Sun Qiong¹, Xie Chongfeng¹, Zhao Huiling¹, Lu Xushan², Liu Shucheng²

(1.Beijing Research Institute of China Telecom Co., Ltd., Beijing 100035, China;

2.Huawei Technology Co., Ltd., Shenzhen 518129, China)

Abstract: SDN-based IPv6 transition technology was researched. A solution on how to introduce SDN-based IPv6 transition technology in operator network was firstly proposed, including application scenario, requirements, overall architecture, key technologies and major deployment requirement. This technology is quite important for operators to introduce SDN-based IPv6 transition in operator network.

Key words: software defined networking, IPv6, transition technology

1 引言

随着运营商 IPv4 地址的日益紧缺,IPv4 向 IPv6 过渡已成为运营商的必经之路。IPv6 能够从根本上解决 IPv4 地址不足的问题,恢复业务的端到端特性,从而保证业务的持续发展。然而,IPv4 向 IPv6 过渡并不是一蹴而就的。由于 IPv4 与 IPv6 并不兼容,必须引入 IPv6 过渡技术才能实现 IPv4 到 IPv6 的转换,一方面实现过渡期间 IPv4 地址的复用,从而节省 IPv4 公网地址消耗;另一方面实现 IPv4 与 IPv6 的转换及网络的跨越。

虽然过渡技术有很多种,包括隧道类、翻译类等,并且适用于不同的应用场景,但究其实质均为流表的处理问题,实现将 IPv6 地址、端口信息与 IPv4 地址及端口信息的映射,并按照隧道或翻译方式予以实现。该实现思路与

SDN (software defined networking, 软件定义网络) 的流表处理不谋而合。

SDN 采用控制与转发分离的思路,控制器(controller)具有全局视角,统一实现策略的制定、流表的下发等操作,而转发器(switch)接收来自控制器的请求,执行相关流表的操作。SDN 早期是在斯坦福大学提出的 OpenFlow 基础上发展起来的,目前已广泛受到包括思科、Juniper、华为等在内的业界主流厂商和运营商的深度参与,并有 ONF、NFV、ITU-T、IETF、BBF 等多个标准化组织制定相关标准。

虽然 SDN 技术已得到业界的广泛关注,但如何将 SDN 技术与 IPv6 过渡技术相结合,从而解决运营商在 IPv6 过渡时期的复杂性和业务开展问题,目前还没有进行深入研究。本文深入研究了在运营商网络中将 IPv6 与 SDN 相结合的应用场景及实施优势,并首次提出了



Openv6 系统实现架构与关键技术,重点研究了现网部署实施的可行性。

2 基于 SDN 架构的 IPv6 过渡技术设计

2.1 业务需求

在 IPv4 向 IPv6 过渡时期,存在多种复杂的应用场景和过渡技术,运营商对不同的过渡技术进行选择、切换以及管理时都会存在不同的需求。通常,由 IPv4 向 IPv6 的过渡需要经历 IPv4 主导、IPv4/IPv6 双栈以及 IPv6 主导 3 个不同的阶段,在不同的阶段又有多种应用场景并存。因此,在整个 IPv4 向 IPv6 的过渡时期,存在如下业务承载需求。

(1) 不同过渡技术间的切换

在 IPv4 向 IPv6 过渡的不同时期,需根据网络中当前设备的支持情况、网络中 IPv4/IPv6 流量的变化情况、IPv4 剩余地址的数量等多个因素,选择合适的过渡技术。例如,在网络中 IPv6 支持情况较好、终端对过渡技术支持较好时,可以采用基于 IPv6 的过渡技术(如 DS-Lite 等),以更好地适应将来的发展;若网络中当前设备及终端的 IPv6 支持能力较低,则可以采用一些基于双栈的过渡技术(如 NAT444 等),以降低对前期网络中 IPv6 能力的要求。这些技术的选择需要因地制宜,不能一概而论。

(2) 多种过渡技术的共存

在 IPv4 与 IPv6 的同一过渡时期,可能存在多种过渡技术用于不同的应用场景。例如,隧道类过渡技术可解决 IPv4 地址的后向兼容性问题,翻译类过渡技术可以解决 IPv4 与 IPv6 应用互通的问题。当多种过渡技术共存时,对设备的实现提出了一定的要求。当前设备在实现多种技术共存时还存在一定程度上的限制,如多技术叠加时设备资源(如存储空间等)无法灵活共享,需要为不同的技术开辟一个独立的计算与存储空间等。因此,当前设备实现无法最大程度地充分利用设备的硬件资源与计算资源。

(3) 统一地址池的管理

在 IPv6 过渡时期,不同的设备、过渡技术所使用的

IPv4 地址池通常是单独管理的,不同的过渡技术通过不同的实例进行管理和配置。但由于在 IPv6 过渡时期,过渡技术的共存、选择和迁移是随着 IPv6 和 IPv4 的流量变化而变化的,因此,需要对 IPv4 地址池不断进行调配,并根据业务量的变化进行调整。此外,由于运营商现有的 IPv4 地址资源已非常零散,多种过渡技术的引入无疑会进一步加剧地址池的碎片化,使得地址池的管理和维护更加复杂。

(4) 网络设备的虚拟化

在 IPv6 过渡时期,不同的过渡设备在进行相关的流表处理时,都需要消耗大量的计算资源、存储资源等,用于存储用户的映射信息等内容,这些内容会随用户流量、业务状况的变化而变化。网络功能的虚拟化能够更有效地实现过渡资源的集约化管理。

(5) 服务质量的协同控制

随着网络智能化技术的逐步增强,运营商在网络中实现业务服务质量的动态调整、路径优化、业务自动化部署的能力大大提高。然而,随着 IPv6 在现网中的引入,网络中地址转换点的增加使得用户、业务的识别与调整的复杂度进一步增大。

2.2 当前主流过渡技术的模块化实现

当前,多种过渡技术对用户侧设备有不同的要求,如 DS-Lite 为纯 IPv6 接入过渡技术,一方面要求用户侧设备实现隧道的封装等功能,另一方面仅申请 IPv6 地址即可;NAT444 作为双栈接入过渡技术,用户侧设备通常包括 NAT 功能,同时也需要申请 IPv4 和 IPv6 地址;MAP、LAFT6 过渡技术则需要在家域网中包含 NAT 及隧道功能。这些模块在多种过渡技术中是可以复用的。表 1、表 2 分别总结了当前过渡技术在用户侧和网络侧的相关功能模块。

通过多个功能模块的叠加、组合和复用可以实现多种过渡技术,通过控制面的策略下发可以实现多种过渡的切换,从而可以比较灵活地实现不同业务场景的适配。

表 1 当前过渡技术在用户侧的相关功能模块

过渡技术	IPv6 转发	NAT	隧道封装	其他参数的获取	接入类型
NAT444	√	√			双栈
DS-Lite	√		√		IPv6 单栈
LAFT6	√	√	√	√	IPv6 单栈
MAP	√	√	√	√	IPv6 单栈
6rd			√	√	IPv4 接入

表2 当前过渡技术在网络侧的相关功能模块

过渡技术	IPv6 转发	NAT	隧道封装	接入类型
NAT444	√	√		双栈
DS-Lite	√	√	√	IPv6 单栈
LAFT6	√	√		IPv6 单栈
MAP	√		√	IPv6 单栈
6rd			√	IPv4 接入

2.3 基于 SDN 的 IPv6 过渡技术功能架构

引入基于 SDN 的 IPv6 过渡技术后的系统架构如图 1 所示,网络边缘转发设备内置控制代理(control agent)模块,用于转发面与控制面的协议交互。

- 转发面主要实现通用 IPv6 过渡技术的模块化操作,统一实现多种类型过渡技术的转换处理。转发面和控制面的协议交互可以在 OpenFlow 协议基础上进行扩展,也可以在现有协议(如 RADIUS、Netconf 等)基础上进行扩展。
- 控制面主要实现不同过渡技术的下发、流表/规则的下发与控制以及过渡期地址资源的管理。
- 应用层则主要通过与控制面的交互,实现对不同过渡技术的管理与配置。

2.4 功能模块组成

基于 SDN 的 IPv6 网络参考架构包括三大功能模块,分别为:应用功能模块、SDN 控制器相关功能模块和转发节点相关功能模块。

(1)应用功能模块

应用层主要提供 IPv6 演进技术插件、地址池及网络资源管理、服务质量的动态调控及其他第三方业务的开放功能。

其中,应用层为 IPv4 向 IPv6 过渡中的各种技术提供

插件,主要是方便灵活配置网络转发节点,以适应过渡技术的不同场景和不同阶段;地址池及网络资源管理功能为运营商提供地址池管理、资源管理的统一界面,并能够与控制面交互地址的相关信息;服务质量的动态调控与差异化管理功能,为特定用户、业务提供带宽、路径的差异化提供能力;第三方业务开放接口为第三方提供业务管理能力。此外,应用层还可提供综合的业务运营和管理功能,包括网络管理、系统管理、计费、营业、账务和客户服务等。

(2)SDN 控制器相关功能模块

控制器层主要包括控制协议功能、过渡技术编排、映射表项下发、对网络资源的管理、网络虚拟化以及基础协议相关功能。

其中,控制协议功能提供对转发设备的控制协议(如 OpenFlow 等),与转发节点建立控制协议连接,其他功能模块可以通过控制协议对转发节点进行控制;过渡技术的编排功能支持对指定的过渡技术进行编排,下发给指定的转发设备,并进行相应的配置管理;网络资源管理功能支持对映射表项、设备资源占用情况的统计查询,并可基于当前资源的使用情况确定下发的转发设备列表,以实现负载均衡和冗余备份,且能在资源不足时动态调整路径,并下发给相应的转发器,同时支持针对各种 IPv6 过渡技术,对地址资源(包括地址和端口)进行统一管理,包括地址的资源分配、地址资源回收、地址资源再分配等;网络虚拟化功能能够实现对网络的抽象,屏蔽物理网络细节。

此外,控制器还可以更细粒度地控制转发面映射表,映射表项的下发支持基于转发层上送的首个分组及过渡策略来确定相应的映射表项,并下发给转发设备;支持对特定映射表项的生成、更改与删除。同时,网络协议层还可

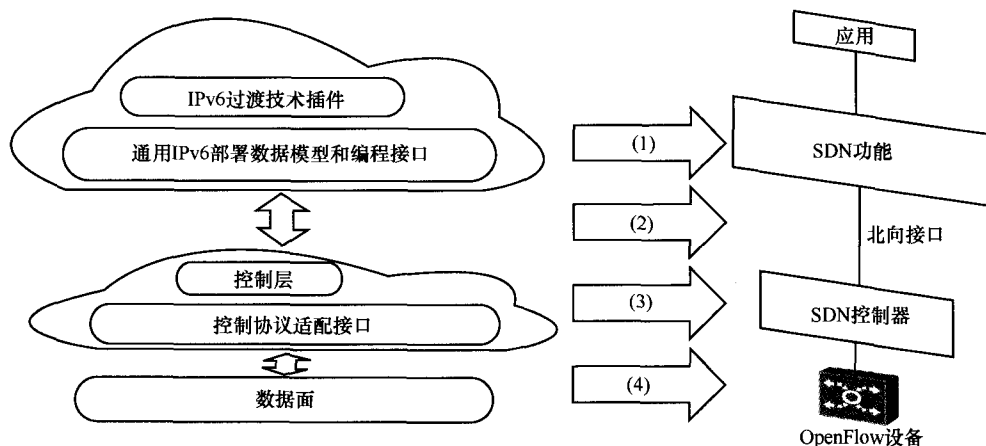


图1 基于 SDN 的 IPv6 过渡体系组网架构



提供传统路由器具备的路由协议栈,使 SDN 控制器能够代替转发节点对外运行传统路由协议。

(3)网络侧转发节点功能模块

网络侧转发节点功能模块主要包括基础路由功能、控制协议、控制代理、流转发表及其他支撑模块。

其中,基础路由功能运行 IGP 路由协议,负责将基本路由打通,在控制器能够控制转发节点之前,建立控制通道以及网络拓扑收集;控制协议通过与控制器的控制协议层对应,负责与控制器建立和维持控制协议连接,如 SNMP、OpenFlow 等协议;控制代理负责向 SDN 控制器注册、上报信息,接收 SDN 控制器下发的控制信息和转发表项,并下发到转发面以指导转发;流转发面可实现对报文的转发和处理,实现 IPv6 过渡映射表的统一存储和处理。

(4)用户侧转发节点功能模块

用户侧转发节点功能模块主要包括基本转发及模块化实现、基本模块的编排、控制代理和监测模块。

其中,基本转发及模块化功能负责基本的 IPv6 转发、不同基本模块的实现(包含 NAT、隧道等基本单元模块);基本模块的编排可以支持不同的模块化组合,并可以接受不同过渡技术的模式及基于特定模块的编排;控制代理负责向控制器注册、上报信息;监测模块负责对本地转发情况、地址端口占用情况进行监控。

3 关键技术

3.1 流表选择性映射技术

在 IPv6 过渡技术中,面向单流的过渡技术(如 DS-Lite、NAT444、NAT64 等)是在流表到达转发器后才发起新建流表的请求,这对于转发器和控制器之间的设备接口、带宽会有很大的压力。对于这类技术,应用 SDN 技术的主要优势在于,可以较为便捷地实现状态信息的设备间同步。而由于状态信息的同步对于短连接的业务和长连接的业务会有较大的差异,短连接的业务在连接重新建立后可以继续原有的业务体验,而长连接的业务在连接重新建立后会中断原有的业务体验。因此,流表选择性映射技术可以选择配置长连接的业务,将这类业务的映射流量在多个转发器之间进行同步。

流表选择性映射机制的实现如图 2 所示,数据流的实现过程介绍如下:

(1)由编排层向控制器通告需要进行状态同步的目的地址,并由控制器将该请求下发给转发器;

(2)数据流到达转发器;

(3)转发器查找本地状态同步映射表,若找到,转发器向控制器请求获取流表;

(4)控制器向转发器备份组统一发送数据流;

(5)数据流经过转发器转发。

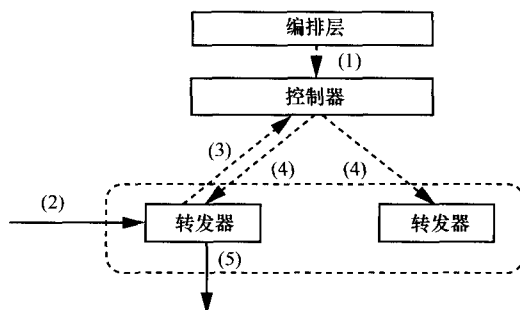


图2 流表选择性映射机制的实现

由控制器下发给转发器的状态同步表项的 ACL 匹配规则可以基于特定的目的地址,也可以基于一定的协议(如 TCP 及端口等),在收到协议时向控制器发起新建流表的请求。在该方案中,如果控制器发生宕机,取消由控制器下发的 ACL 匹配规则,从而可以实现由 SDN 转发器向传统转发器的转换。

该方案不仅可以大大降低基于每流状态的 DS-Lite/NAT444 方案中控制器的处理压力,同时也适用于与现有过渡设备的共存,当控制器发生宕机时不会受到较大的影响。

3.2 地址池自动化管理技术

当一个控制器同时管理多个转发器时,多个转发器的地址池可以采用虚拟化管理的方式实现,即将一个地址池划分为多个子地址池单元,控制器在收到一个关于新建状态表项的请求时,基于不同的过渡技术在不同的地址池单元中进行分配,地址池单元可以不连续排列,从而提高地址池的整体利用率,如图 3 所示。

3.3 设备状态组管理

为了实现转发器间的负载均衡和冗余备份,需要在控制器中建立转发器的负载均衡组和冗余备份组。其中,位于负载均衡组内的转换设备具有近乎相同的状态表数量,状态信息不在多个设备间共有;位于冗余备份组内的转换设备同时保存相同的状态表,如图 4 所示。

4 试验与验证

为了验证该架构的性能,在真实环境中进行测试,测试环境如下:

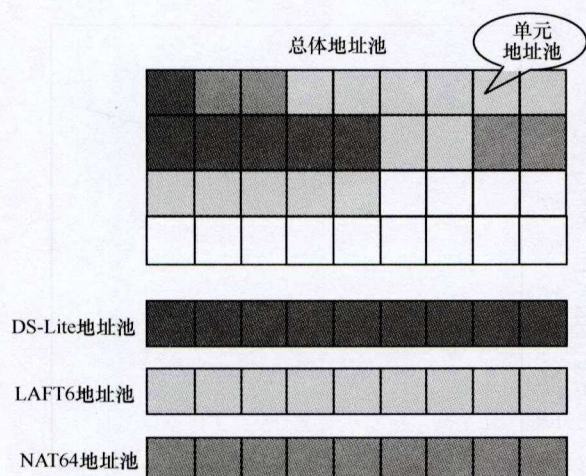


图3 地址池自动化管理技术示意

- 在运行 Open vSwitch 的普通商用硬件上实现基于 SDN 架构的 IPv6 过渡技术；
- IPv6 过渡技术的应用运行在配置为 Intel Xeon E5-2407 CPU、32 GB 内存的计算机上；
- 数据面运行在配置为 Intel Core i5-2400 CPU、4 GB 内存的计算机上；
- 所有硬件平台上安装 64 bit CentOS 作为操作系统。

该系统部署在一个约有 1 000 人使用的网络中，系统架构如图 5 所示。

收集了一周工作时间时的数据，数据呈现如图 6、图 7 所示。图 6 展示了对于用户体验的重要参数——流建立时间的验证，流建立时间指用户建立一个新的会话需要消耗

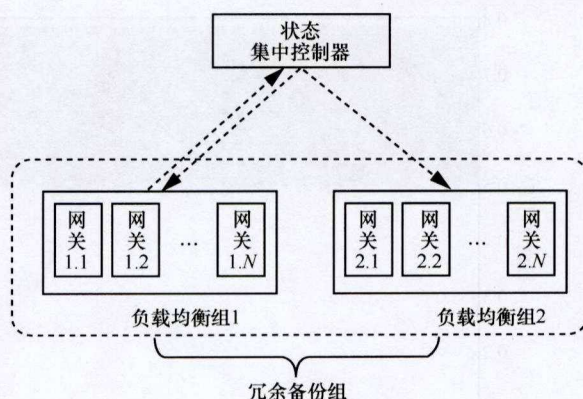


图4 设备状态组管理示意

的时间，如用户打开一个新网页所需的时间。图 6(a)展示了在不同的时间段收集到的流建立时间的数据，图 6(b)展示了该数据的统计结果，即流建立时间小于 0.7 ms 的百分比。可以看出，该架构的流建立时间是可商用、可部署的。

类似的，图 7 展示了另一个重要参数——数据面的分组时延，这个参数可以反映用户真实的体验，如 VoIP 的时延情况。图 7(a)展示了在不同时间段收集到的数据面的分组时延的数据，图 7(b)展示了该数据的统计结果，即数据面的分组时延小于 4 ms 的百分比。可以看出，该架构的数据面分组时延同样是可商用、可部署的。

5 部署与实施考虑

基于 SDN 的 IPv6 过渡技术在部署时要充分考虑与现有网络设备的兼容性，如图 8 所示。IPv6 过渡技术的流表可以

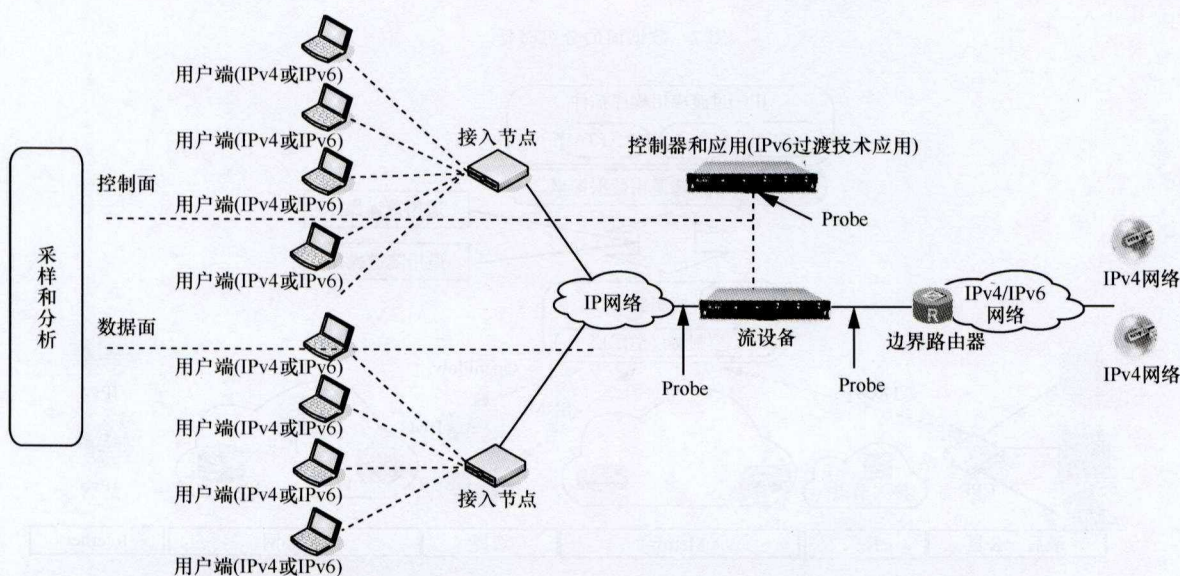


图5 真实环境中的系统架构

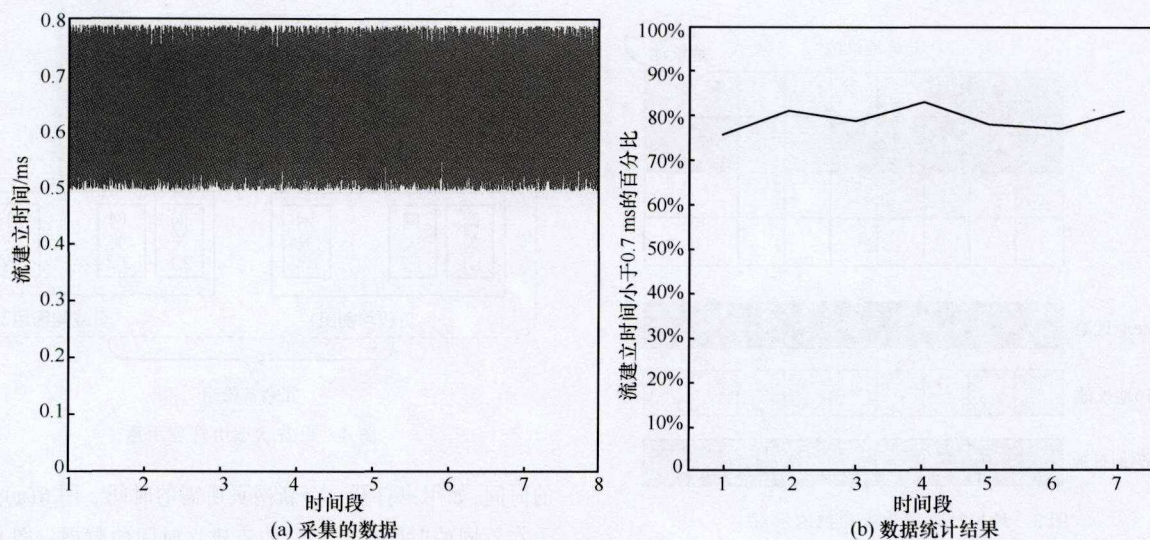


图6 流建立时间

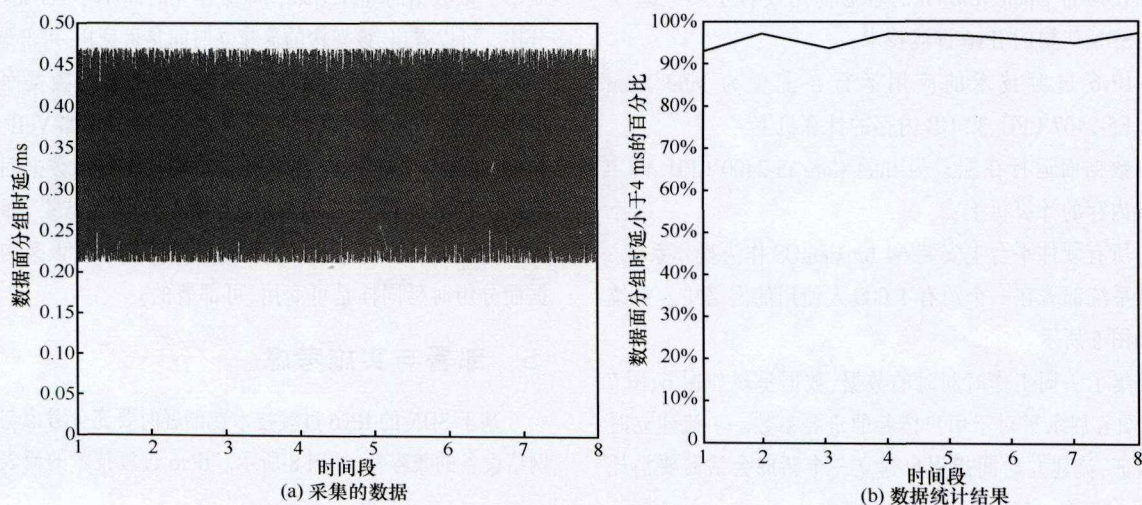


图7 数据面的分组时延

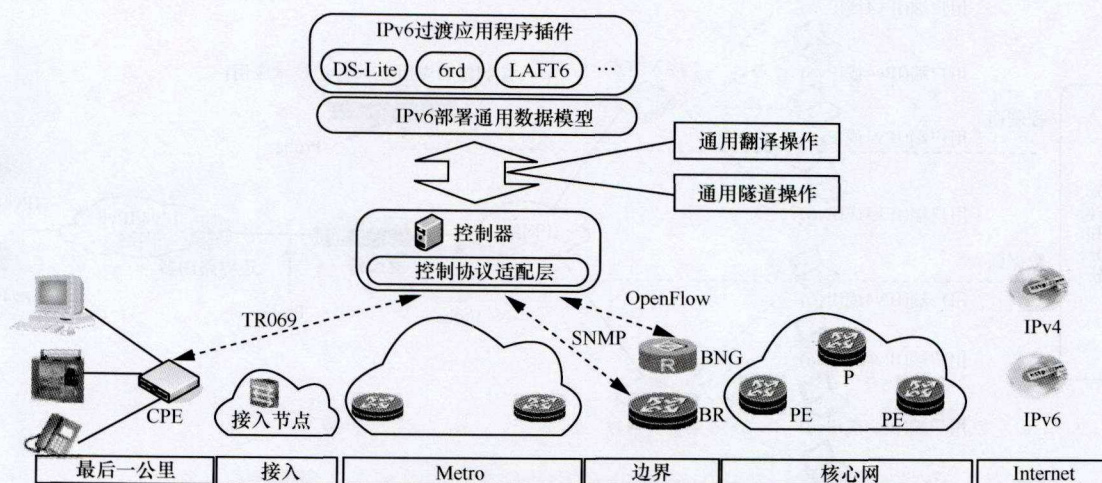


图8 基于SDN的IPv6过渡技术的部署

本地生成,也可以采用首个分组上送的方式由控制器生成。

在实际部署时,考虑到 IPv6 过渡设备会在 SDN 技术完全成熟前进行部署,因此需采用分阶段的部署策略,具体介绍如下。

第一阶段,现网中已部署部分 IPv6 过渡设备,此时基于 SDN 的 IPv6 过渡设备可以首先在网络侧新增设备中实现。在过渡技术的选择上,尽可能与已部署的过渡技术保持一致。针对新引入的过渡技术,可以首先在新设备中支持。此外,控制器可采用在同一城域网内集中式部署的模式,增加协议接口适配模块,以实现与现有设备的互通。该阶段由于全网并没有完全支持 SDN 技术,对于某些新业务的开展,只能指定在特定的设备和路径中予以实现。

第二阶段,逐渐在家庭网中引入 SDN 技术,从而可以更好地适配新业务的发展。同时,可逐步升级现网中的已有设备,使得控制器能够完全掌握全网的资源配置情况,同时编排层也能完成与现网支撑系统的配合,且对用户的实际端口及地址的使用情况有了更为清晰的认识。此时可逐渐根据实际需求调整特定过渡技术的选择以及对部署位置、资源进行灵活调配。

第三阶段,全网逐渐全面引入 SDN 技术,为了更好地支撑 IPv6 的发展,在 SDN 技术中也会更多地引入网络智能化、虚拟化等技术,以更好地支持物联网、云计算、移动互联网等应用的发展。

参考文献

- 1 Bagnulo M, Matthews P, Beijnum I. Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers. draft-ietf-behave-v6v4-xlate-stateful-11(Work in Progress), March 30, 2010
- 2 Huang B, Deng H. Prefix NAT: Host based IPv6 Translation. draft-huang-behave-pnat-01 (Work in Progress), February 19, 2010
- 3 Boucadair M, Jacquenet C, Grimault J L, et al. Deploying Dual-Stack Lite (DS-Lite) in IPv6 Network. draft-boucadair-dslite-interco-v4v6-03(Work in Process), February 24, 2010
- 4 Li X, Bao C, Chen M, et al. The CERNET IVI Translation Design and Deployment for the IPv4/IPv6 Coexistence and Transition. draft-xli-behave-ivi-07(Work in Progress), January 6, 2010
- 5 Li X, Bao C, Zhang H. Address-Sharing Stateless Double IVI. draft-xli-behave-divi-01(Work in Process), October 26, 2009
- 6 Bush R. The A+P Approach to the IPv4 Address Shortage. draft-ymbk-aplusp-05, October 27, 2009

[作者简介]



孙琼,女,博士,中国电信股份有限公司北京研究院下一代互联网技术总监,主要研究方向为下一代互联网、SDN、虚拟化等,发表论文 20 余篇,参与中国电信下一代互联网规模部署商用工程项目,曾提出创新型 LAFT6 过渡技术,并被 IETF 接受。



解冲锋,男,博士,中国电信股份有限公司北京研究院下一代互联网产品线总监、教授级高级工程师,长期从事网络架构、下一代互联网以及宽带新业务方面的研究和实践工作,负责中国电信和国家多个重大 IP 科研项目,曾获得中国通信学会科学技术进步奖一等奖 1 项、三等奖 2 项,以第一作者身份提出专利申请 11 项。



赵慧玲,女,中国电信股份有限公司北京研究院总工程师,中国电信股份有限公司云计算研究中心主任,清华大学中国电信下一代互联网技术联合实验室主任、教授级高级工程师,主要从事宽带网络和下一代网络的技术研究以及通信网络发展战略研究工作,主持我国网络标准的研究和制定工作。获国家政府津贴,发表文章近百篇,出版技术专著 12 部;获国家技术发明奖二等奖 1 项,国家科学技术进步奖二等奖 2 项,部级科学技术进步奖一等奖 3 项、二等奖 9 项、三等奖 8 项。



卢绪山,男,华为技术有限公司网络融合技术研究部技术专家,多年从事电信宽带网络领域研究工作,主要研究方向为网络 QoS、软件定义网络/OpenFlow、IPv6 等,曾于 SIGCOMM 国际会议发表研究成果论文。



刘树成,男,博士,华为技术有限公司固定网络研究部研究员及标准代表,主要研究方向为 IPv6、网络功能虚拟化、内容网络。

(收稿日期:2014-03-20)