



专栏：随选网络

随选网络系统架构及关键技术实践

孙颖¹, 林睿², 聂世忠¹

(1. 中国电信股份有限公司北京研究院, 北京 102209;

2. 中国电信集团公司, 北京 100032)

摘要: 分析了中国电信对 SD-WAN 应用场景和技术方案的选择, 同时, 对中国电信随选网络系统整体架构和研发涉及的关键技术进行了阐述, 包括 EVPN 技术、Devops 平台技术、VxLAN 中继引流技术的选择和应用方案。

关键词: 随选网络; SD-WAN; 编排器; 控制器

中图分类号: TN929

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017323

System architecture and key technologies of on-demand network

SUN Ying¹, LIN Rui², Nie Shizhong¹

1.Beijing Research Institute of China Telecom Co., Ltd., Beijing 102209, China

2.China Telecom Corporation, Beijing 100032, China

Abstract: China Telecom's selection of SD-WAN application scenarios and technical solutions was analyzed. At the same time, system architecture of China Telecom's on-demand network and key technologies involved in R&D were elaborated, including EVPN technology, Devops platform technology, selection of VxLAN relay drainage technology and application.

Key words: on-demand network, software defined wide area network, orchestrator, controller

1 引言

中国电信作为国内最大的基础网络服务商之一, 通过 SDN/NFV/云计算技术强化大网优势、提升业务体验是首要着眼点。因此中国电信随选网络实践中, 选择了以软件定义方式建立端到端 WAN 路径为业务切入点, 结合质量、可靠性等需求, 为客户提供分钟级业务开通时限以及后续业务使用中的自动化调整功能。在技术实现方面,

选择了对现网改造要求最小的 overlay 方案, 作为现网实施的第一步, 即基于 SDN 技术在基础网络之上实现自动化、智能化的端到端逻辑通路, 对客户屏蔽底层网络的复杂性。这条逻辑通路可以构建在互联网上, 也可以构建在 CN2 网络(中国电信高等级骨干网)上, 对客户接入方式没有特定要求。

从 2006 年 SDN (software defined network, 软件定义网络) 概念诞生至今, 国内外业界一直

收稿日期: 2017-11-10; 修回日期: 2017-12-07

在探索和扩大其应用领域。SDN 初期关注点在云和数据中心内,逐步延展到互联网、广域 IP 网、传输网,近两年聚焦于企业市场和广域网范畴的 SD-WAN (software defined wide area network, 软件定义广域网)。所谓 SD-WAN,既有技术内涵,又有业务内涵。技术内涵,即 SD-WAN 是以 SDN 技术为基础,实现广域网通道开通、传送质量改善、客户成本降低等相关技术;业务内涵,即 SD-WAN 基于以上技术向客户提供 WAN 服务。AT&T、NTT、DT 等全球多家运营商以及众多新兴服务商在 SD-WAN 服务中,都选择了面向政企客户、以企业互联和云/数据中心互联为主的业务场景。在主要业务功能上,可以归纳为 3 类:IP VPN (virtual private network, 虚拟专用网络) 业务的自助订购和敏捷发放、入云/云间通道的自动配置、专网与互联网混合路径选择。在这三大类业务功能中,又有多种技术实现方式,各方均依据自身的既有技术优势实现,无分优劣,各有千秋。笔者认为,以上这些业务形态均属于 SD-WAN 范畴,随选网络即中国电信在 SD-WAN 领域的一次深入实践。

从技术路线上来说,中国电信的随选网络是基于 WAN 中的集中控制,结合质量、可靠性等需求,以软件定义的方式建立端到端 WAN 逻辑路径,提供日常业务调整和查看手段。

2 系统架构设计

中国电信随选网络系统基于 SD-WAN 技术路线,以 SDN 技术为基础,整个系统由业务门户、SDN 编排器、SDN 控制器、SDN 设备四大组件组成,如图 1 所示。

图 1 中,业务门户是业务自助订购与网络可视化操作界面,提供客户角色、客户经理角色、网络管理员角色的相关功能;SDN 编排器北向提供面向业务的网络能力封装和业务功能编排,南向整合多厂商控制器接口,实现跨厂商、跨业务的

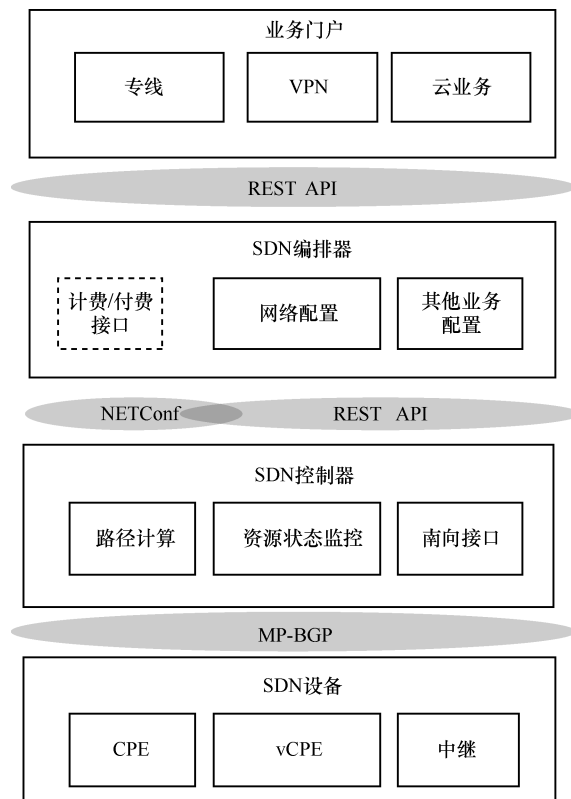


图1 随选网络系统四大组件

能力统一化;SDN 控制器实现对底层 SDN 设备的配置下发、资源状态收集和路径计算以及异常情况下的设备倒换;SDN 设备执行 SDN 控制器配置指令、完成相关数据处理和转发。在中国电信随选网络方案中,SDN 设备主要包括 CPE (customer premise equipment, 客户终端设备)、vCPE (virtual customer premise equipment, 虚拟客户终端设备)、VxLAN (virtual extensible LAN, 可扩展虚拟局域网) 中继。

应该指出的是,在业界诸多具体实现中对 SDN 控制器、编排器的功能界定各异,有的厂商甚至将编排器和控制器合一。参考图 2 的模型,可以看到在整体架构中,控制器北向提供功能模型,编排器北向提供业务模型。在一些具体实现中,控制器可以实现业务模型,由控制器直接对接业务平台或门户,这种做法虽然在业务场景单一情况下实施简便,但未来会对异厂商互通及编排器的业务功能编排造成障碍。

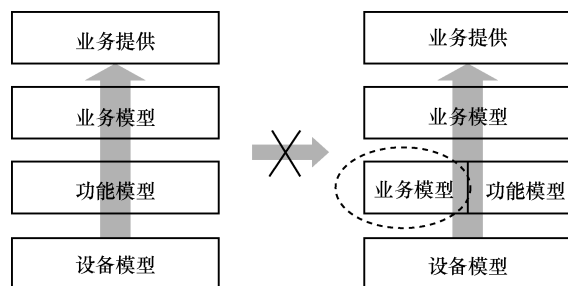


图2 SDN 系统层次划分

SDN 编排器决定业务能力和业务形态，为业务运营单位提供业务编排手段，决定业务编排操作难易度。对编排器来说，业务是一系列相关事务，事务对应一些串行的配置动作，这些动作全部有序配置成功即完成一个业务功能，有一个配置动作没有成功即视为业务配置失败，其他配置动作全部回滚。因此，从系统设计角度来说，编排器是对业务进行理解的“大脑”，真正被上层软件定义的关键就在编排器这一层。在随选网络的架构设计中，要求编排器和控制器严格按照业务模型和功能模型进行划分，控制器不做业务层面的组织工作，只向编排器提供单纯的、细颗粒度的原子能力，这样有利于日后编排器根据业务需求进行灵活的业务功能编排。

从部署方式上说，编排器和控制器可部署在云资源池中，相互 IP 地址可达即可。由于编排器开发采用了松耦合微服务方式，因此编排器各软件模块及原子能力可根据需求分布式部署，相互 IP 地址可达即可。编排器作为系统资源编排的统一入口，建议全国或全省一套；而控制器通常是单域、单专业的。从编排器视角看，对控制器的部署方式没有特别要求。编排器和控制器的作用主要是控制面的配置下发，并不涉及转发面的处理，因而一般不会出现高并发重负载的情况。

3 系统关键技术

3.1 EVPN 技术

随选网络系统控制器和设备开发采用 EVPN

(ethernet VPN, 以太网 VPN) 技术，控制面为 MP-BGP，数据面为 VxLAN。EVPN 技术体系如图 3 所示。

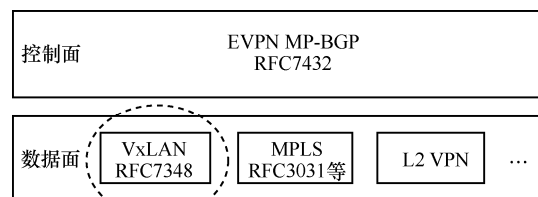


图3 EVPN 技术体系

MP-BGP 在 BGP 基础上扩展了 RD (route distinguisher, 路由区分符) 和 RT (route target, 路由目标符) 两个标识，实现 VPN 路由隔离。在中国电信随选网络系统四大组件中，要求 SDN 控制器和 CPE/vCPE 设备支持 MP-BGP，按需建立 BGP 邻居关系，从而实现跨广域的二/三层路由信息交换。要求 CPE/vCPE 支持 VxLAN，作为 VxLAN 的 VTEP (VxLAN tunnel end point, VxLAN 隧道端点)，提供 VxLAN 隧道数据分组封装和解封装功能。实施初期，由于网络规模小，节点数量少，可采用纯分布式路由方式，SDN 设备运行 MP-BGP，完成路由信息交互。后续随着网络规模扩展，计划在控制器层面增加 RR (路由反射器) 模块，与所有节点建立邻居关系，解决大量节点之间 Full Mash 的 N^2 问题。由于采用了 EVPN 技术，中国电信随选网络系统可同时提供二/三层 VPN 以及 Full Mash 和 Hub-Spoke 组网方式。在 CPE/vCPE 处理 VxLAN 封装和解封装时，需识别内层源 IP 地址与目的源 IP 地址是否属于同一子网，如属同一子网，则启用二层 VPN 处理能力，即桥接方式，如图 4 所示；如属不同子网，则启用三层 VPN 处理能力，即路由方式，如图 5 所示。

随选网络系统选用 EVPN 技术，还基于以下几点考虑。

- 控制面采用成熟的、标准化的 MP-BGP，更适用于运营商的网络规模和业务稳定性

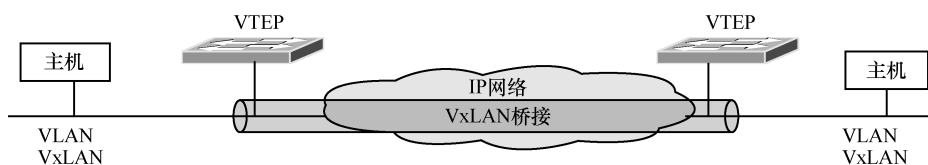


图4 VxLAN 桥接方式

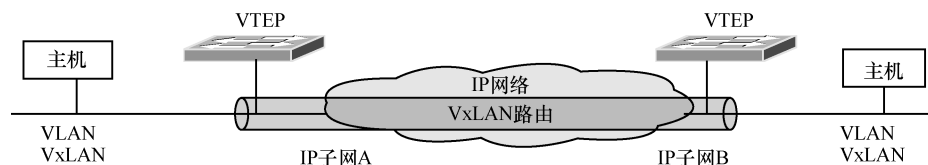


图5 VxLAN 路由方式

要求。

- 通过 MP-BGP 控制 MAC/ARP 的学习，而非 L2 广播自学习方式，有效避免广播风暴，保障业务安全性。
- 通过 MP-BGP 实现 VxLAN 隧道的自动创建，简化业务配置，降低网络部署难度。

3.2 Devops 平台技术

随选网络系统中 SDN 编排器的开发引入了 DevOps 平台及其理念，通过内部流程与自动化工具的整合，将开发、测试、部署、运维构成自动化闭环，如图 6 所示。

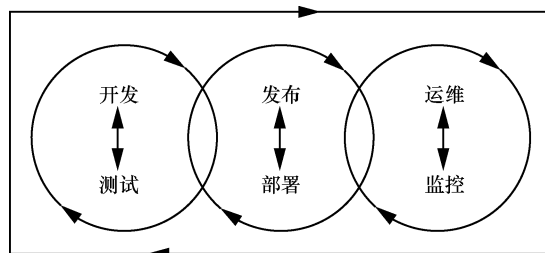


图6 Devops 自动化闭环

DevOps 流程中的集成与测试、部署、运维 3 部分覆盖了从软件产品开发到运维的全过程，通过管理工具（例如 Jenkins）完成 3 部分间的无缝衔接。代码被提交到源代码管理服务器上，自动触发代码的构建与单元测试，在单元测试成功后触发功能测试和压力测试。测试的范围不仅包含软件代码部分，也包含编排形成的业务模型、策略

模型等。在开发中选用 Git 作为源代码管理工具，Jenkins 作为持续集成与部署的工具，Ansible 作为配置管理工具，Docker 和 Kubernetes 作为部署工具。

中国电信随选网络编排器基于 DevOps 平台开发，架构上分为开发态和运行态，软件从开发态送入沙箱调测，测试通过自动进入正式系统运行，整个流程中无需人工干预。软件模块之间采用通用总线进行调用、寻址、负载均衡等。同一软件模块可以部署在不同的物理位置、不同实例上。

从软件系统架构的角度看，随选网络系统 SDN 编排器的目标既包含了稳定可靠的软件系统，也包含了用于实现新需求新功能的集成开发环境。基于这套自动化测试部署的底层平台，还将实现面向网络维护人员的业务编排界面，运维人员只需理解业务逻辑，通过简单的拖拽动作就可以完成新的业务能力配置，不需要学习命令和代码技术。因此，这套系统也为企业实施新型人力资源组织提供了技术手段，要求未来的业务运营团队既包含研发测试团队，也包含传统运维团队，两部分人员有机融合。

3.3 VxLAN 中继引流技术

在一些业务场景中，希望客户的流量能自动进入 CN2 网络中，从而规避可能的网络拥塞点，实现差异化承载。这个功能是通过随选网络系统



中的 VxLAN 中继来实现的。

VxLAN 中继引流实现原理如图 7 所示。

中国电信随选网络系统的 VxLAN 中继设备部署于引流目标网络的边缘，负责前段 VxLAN 隧道的终结和后段通道（可以是 VxLAN 隧道，也可以是 MPLS 通道）的数据分组封装。客户驻地网出口位置的 CPE 通过识别数据分组的目的 IP 地址，将需要导流的数据分组分离出来，在封装 VxLAN 数据分组时，将外层目的 IP 地址填上 VxLAN 中继的 IP 地址。这样，这部分数据分组就会沿着引流策略设置好的通道进行传送。

当数据分组离开引流目标网络时，还需经过一个 VxLAN 中继，终结目标网络中的数据通道，重新封装一个 VxLAN 数据分组，其外层目的 IP 地址是对端 CPE 的 IP 地址，从而将数据分组继续向目的地进行传送。可见，整个传送通道是由 3 段通道粘接而成的，两端 CPE 和 VxLAN 中继之间的粘接关系由 SDN 控制器进行配置。

通过 VxLAN 中继引流方案解决客户高等级业务的承载问题，好处在于客户侧不必有两条接入链路（分别接入专网和互联网），对客户的接入方式也没有特殊要求。客户流量经引流绕

过网络中最易发生拥堵的链路段，能够明显提升传送质量。

另外一个实现引流的方案是预先配置城域网 VPN，客户数据在 BRAS（broadband remote access server，宽带远程接入服务器）设备进行分流，如图 8 所示。

中国电信各城域网均双挂 163 和 CN2 网络，可通过策略路由配置的方式使得城域网 VPN 流量进入 CN2 MPLS VPN，这样一条端到端 VxLAN VPN 共享的基础网专用通道就搭建完成了。在面向客户的业务开通层面，SDN 编排器和控制器根据相关业务参数，向分别部署在客户网络和对端系统中的 CPE/vCPE 设备下发参数，业务发放即可在几分钟之内完成。

目前，随选网络系统已经在江苏、浙江、重庆、广东、江西等多个省份完成部署，以统一架构实现了随选园区、随选云网协同、物联网随选专网、随选 VPN4 个业务场景。

4 结束语

在随选网络系统设计中，更多地考虑了系统面向业务的可扩展性和开放性，业务需求能够通

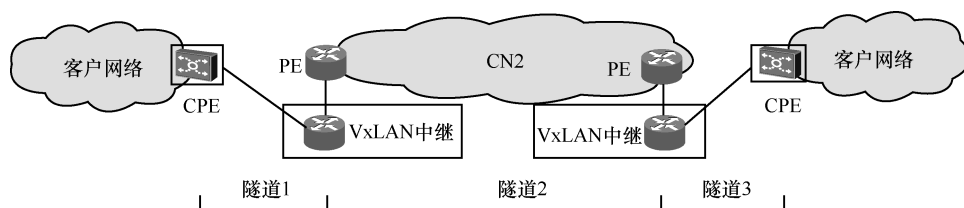


图 7 VxLAN 中继引流实现原理

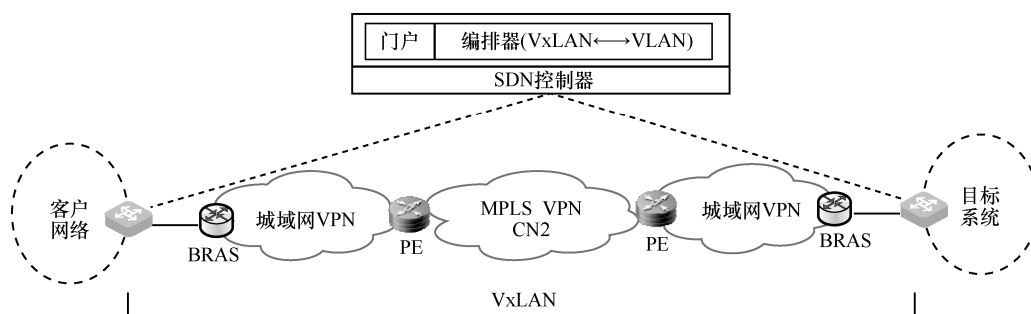


图 8 城域网 VPN 方式引流实现原理

过软件实现来定义网络, 通过网络功能原子化, 使得多种网络能力能够在业务层面进行整合、统一呈现。同时面向应用开发者和服务提供商, 提供便捷的应用开发环境及业务交付的服务支撑。在承载技术层面, 随选网络系统充分考虑了 overlay 技术与 underlay 网络资源相结合的方式, 可以充分体现运营商的差异化竞争力。因此, 随选网络系统的架构设计和关键技术的选择, 能够较好地适应当前业务创新需求以及后续网络演进需要。

参考文献:

- [1] 韦乐平. SDN 的战略性思考[J]. 电信科学, 2015, 31(1): 7-12.
WEI L P. Strategic thinking on SDN [J]. Telecommunications Science, 2015, 31(1): 7-12.
- [2] 何晓明, 冀晖, 毛东峰, 等. 电信 IP 网向 SDN 演进的探讨[J]. 电信科学, 2014, 30(6): 131-137.
HE X M, JI H, MAO D F, et al. Discussion of evolution of carrier IP network to SDN[J]. Telecommunications Science, 2014, 30(6): 131-137.
- [3] 赵慧玲, 史凡. SDN/NFV 的发展与挑战[J]. 电信科学, 2014, 30(8): 13-18.
ZHAO H L, SHI F. Development and challenge of SDN/NFV[J]. Telecommunications Science, 2014, 30(8): 13-18.
- [4] IETF. BGP MPLS-based Ethernet VPN: RFC7432[S]. 2015.
- [5] IETF. Virtual extensible local area network (VxLAN): a framework: RFC7348[S]. 2014.
- [6] 李晨, 陈俏钢, 李凤凯, 等. SDN 的网络模型及北向接口[J]. 中兴通讯技术, 2016, 22(6): 17-21.

LI C, CHEN Q G, LI F K, et al. Network model and northbound interface of SDN[J]. ZTE Technology Journal, 2016, 22(6): 17-21.

- [7] 徐洪磊, 孙琼. 基于 SDN 的网络调优及控制器北向接口探讨[J]. 电信技术, 2016(6): 49-52.

XU H L, SUN Q. Network tuning and discussion on northbound interface of controller based on SDN[J]. Telecommunications Technology, 2016(6): 49-52.

[作者简介]



孙颖 (1977-), 女, 中国电信股份有限公司北京研究院高级工程师、随选网络解决方案专家, 主要研究方向为 IP 网络技术、SDN 系统研发, 拥有多项技术专利和国际标准, 曾获得中国电信集团公司科学技术进步奖二等奖。



林睿 (1975-), 男, 中国电信集团公司网络发展部网络演进和规划处处长, 主要研究方向为承载网架构、SDN 技术、业务平台技术等, 多次获得中国电信集团公司科学技术进步奖一、二等奖。



聂世忠 (1973-), 男, 中国电信股份有限公司北京研究院高级工程师、SDN 技术专家, 主要研究方向为 IP 网络技术、IPRAN 技术、SDN 系统研发, 拥有多项技术专利和国际标准, 曾获得中国电信集团公司科学技术进步奖一等奖。