



vCPE 中 CPE 与云平台连接建立方法

扶奉超, 谢元宝

(中国电信股份有限公司广州研究院, 广东 广州 510630)

摘要: SDN(软件定义网络)和 NFV(网络功能虚拟化)技术已成为当前的研究热点。vCPE(虚拟用户驻地设备)的主要思路是将原来的 CPE 一分为二,保留原有用户侧 CPE 的二层功能,而将三层功能以及业务处理功能上移至云平台中。采用此方法,不仅可以简化原有 CPE 实体网关设备,降低用户操作维护的难度,同时也可降低上门服务概率,并为用户提供增值服务。研究了 vCPE 中用户侧 CPE 如何与网络侧云平台建立连接的问题。通过采用 VxLAN 和 NvGRE 技术,在实现用户侧 CPE 与云平台建立二层通信的同时,对现有海量 CPE 的改动较小。

关键词: 宽带接入网;CPE;SDN;NFV;VxLAN;NvGRE

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2016285

Method of connection establishment between CPE and the cloud platform

FU Fengchao, XIE Yuanbao

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

Abstract: SDN and NFV technology have aroused more and more attention. The main idea of vCPE is to split the CPE into two parts and to make the L3 functions up to the network. Through this method, CPE can be simplified and the difficulty for subscribers to operate and maintain CPE can be reduced. For operators, vCPE can reduce the probability of on-site service, and can provide value-added services for subscribers. The problem of connection between the CPE and the cloud platform was studied. By using VxLAN and NvGRE, the L2 connection between CPE and cloud platform could be established, and at the same time, the change to existing huge CPEs was small.

Key words: broadband access network, customer premise equipment, software defined networking, network function virtualization, VxLAN, NvGRE

1 引言

SDN 和 NFV 技术已成为业内讨论和研究的热点。SDN 针对的是网络问题,其核心思想是将网络的控制平面和转发平面相分离,并开放网络的北向接口,实现网络可编程^[1];NFV 技术是将网络设备虚拟化,用通用 x86 服务器实现传统的专用电信设备功能,实现软硬件解耦,从而降低设备

成本^[2]。SDN 技术和 NFV 技术的快速发展为接入网终端设备 CPE(customer premise equipment)的虚拟化提供了基础。

CPE 也称为家庭网关,是家庭网络与外部网络的接口单元,处于接入网的边缘。CPE 的功能是为家庭用户提供三重业务,即数据、语音和视频业务。运营商通过 NAT(网络地址转换)为家庭网络提供 IP 地址,而在其后的家庭内网对运营商并不可见。当家庭网络出现问题时,运营商有



时需要提供上门到户的技术服务,可能造成维修成本高的问题。然而在很多情况下,网络的问题可能是内网的问题,与运营商提供的网络无关。因此开放家庭网络给运营商,方便运营商进行远程问题诊断与解决,可有效降低 CPE 的上门运维成本。此外,开放内网也可以让运营商为用户提供更多增值业务,如家长式上网管理等。

vCPE 可以解决上述问题。vCPE 的核心思想是保留 CPE 原有的无线接入和交换等二层功能,而用云平台的基础设施实现三层功能和业务处理功能,也就是说把三层的功能上收至网络侧。原有的 CPE 被分为家庭网络侧部分和云平台部分。本文把家庭网络侧保留桥接功能的实体叫做 CPE-B,把 CPE 中位于云平台侧的具有三层网关功能的部分称为 CPE-G。云平台的 CPE-G 采用 NFV 技术,用通用的 x86 服务器搭建虚拟化平台,为每个用户按需分配资源。然而实现 CPE-B 和 CPE-G 的连接是个难题。一方面,为了实现 NAT 等三层功能,CPE-B 与 CPE-G 最好处于同一个 VLAN。一般情况下,CPE-B 连接的家庭网络是个二层局域网,CPE-G 是位于城域网或骨干网的云端的二层网络。CPE-B 连接至 CPE-G 的接入网络可能要跨越运营商的三层网络。在这样的情况下,CPE-B 如何与 CPE-G 建立二层通信是个难题。另一方面,引入 vCPE 后,为了 CPE-B 与 CPE-G 建立通信,可能需要对原有海量 CPE 进行改造,工程量巨大,实现难度极高。因此如何在利旧现有的海量 CPE 设备的基础上,实现 CPE-B 与 CPE-G 的通信也是个难题。

业内在 CPE-B 与 CPE-G 如何连接建立二层连接的问题上鲜有研究。就目前所能够获得的研究成果中,仅 BBF (Broadband Forum) 的 TR-315 标准^[3]从另外的角度讨论了 CPE-B 与 CPE-G 连接建立的问题。与本文侧重点不一样,TR-315 标准研究的是采用 overlay 技术连接 CPE-B 与 CPE-G 时,AAA 服务器和 BNG (broadband network gateway) 等需要满足的要求。而本文侧重于如何具体用 overlay 技术建立 CPE-B 与 CPE-G 的连接,同时寻求合适的隧道端点以利旧已有 CPE。

本文主要针对 CPE-B 和 CPE-G 的连接问题展开研究。通过引入 VxLAN 和 NvGRE (network virtualization using generic routing encapsulation) 隧道技术,并探索合适的隧道端点,实现在利旧已有 CPE 的基础下,建立 CPE-B 和 CPE-G 的二层连接。此外本文总结了 vCPE 的标准化进展和研究进度。

2 vCPE 接入网模型

CPE 在接入网中的位置如图 1 所示,其中 ITMS+(增强型终端综合管理系统)是 CPE 业务的部署和控制平台,也是 CPE 设备的远程管理平台。通过 TR-069 协议,ITMS+既可以实现对 CPE 上的数据、语音、视频等功能进行部署、管理和控制;也可以实现对 CPE 的远程状态管理、配置、故障管理和软件升级等^[4]。EMS(网元管理系统)负责 CPE 的无源光纤网络(PON)接口物理层和链路层的配置和管理。

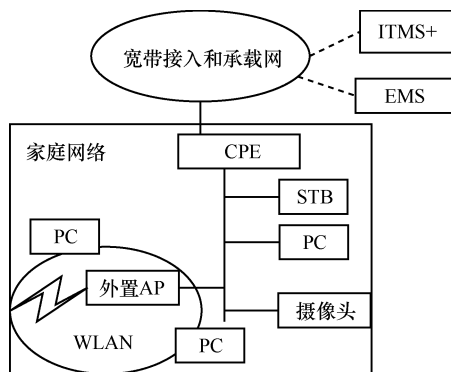


图 1 CPE 在接入网中的位置

采用虚拟化后,vCPE 的网络拓扑如图 2 所示。通过把防火墙、NAT、DPI(deep packet inspection,深度分组检测)、DHCP、DNS 等三层功能转移到网络侧的 CPE-G,简化原有 CPE 网关实体的功能,并使得 CPE-B 支持即插即用的功能。CPE-B 仅保留原有二层的 Ethernet 端口、VLAN 端口、Wi-Fi 端口和 DHCP 客户端等功能,特殊情况下,也可能保留少部分三层功能。

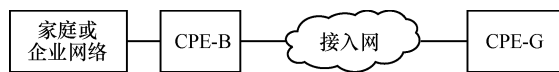


图 2 vCPE 示意

3 CPE-B 与 CPE-G 连接方法

在本文中,基于 overlay 的隧道技术是指利用 IP 网络为 underlay 网络,并以二层网络为 overlay 网络建立隧道连接。vCPE 如何跨越 IP 网络建立二层连接的问题可以通过 overlay 的方法解决,如图 3 所示。采用此种方法可以利用 IP 网络的灵活性。VxLAN 技术和 NvGRE 技术是目前虚拟网络中应用最多的隧道技术,有着良好的应用前景。目前

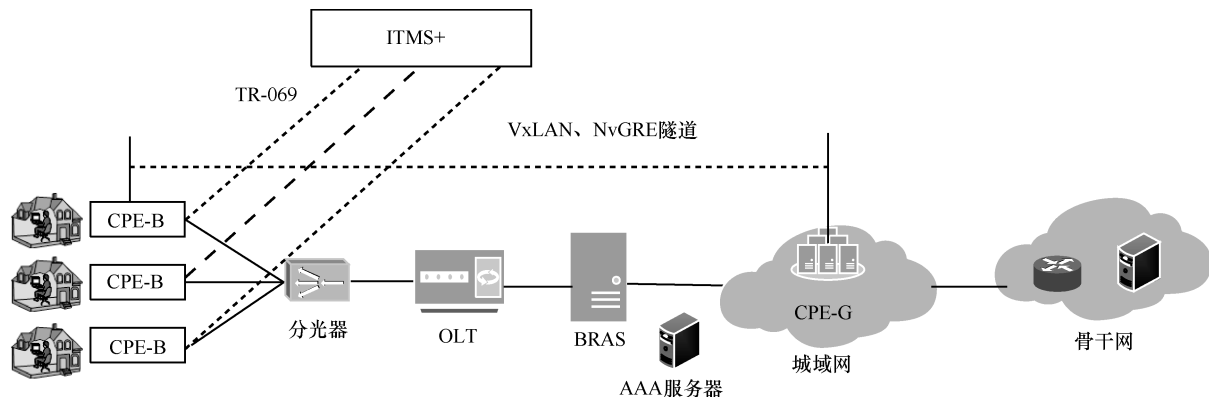


图 3 CPE-B 与 CPE-G 连接示意

在宽带接入网技术方面, PON 不仅具有高容量的优势, 而且具有高可靠性。近年来 PON 技术得到了迅猛发展, 是未来最为看好的宽带接入网技术^[5]。鉴于以上原因, 本文的接入网均以 PON 为例。接下来详述用于 CPE-B 与 CPE-G 连接的 VxLAN 技术和 NvGRE 技术。

3.1 VxLAN 技术在 vCPE 中的应用

VxLAN 不仅可以解决 VLAN 号 12 bit 的限制问题, 还可以利用 IP 网络固有的灵活特性^[6]。应用 VxLAN 建立隧道的方法比较简单, 对基础网络改动较小, 因此本文采用 VxLAN 技术建立 CPE-B 至 CPE-G 的二层通信。

3.1.1 VxLAN 技术的实现过程

VxLAN 是一种 L2 over UDP 的隧道技术。与一般的隧道技术一样, 都需要两端有隧道的端点。在 VxLAN 中, 隧道的端点称为 VTEP (VxLAN tunneling end point, VxLAN 隧道终端), 其中 CPE-B 侧的 VTEP 记为 B-VTEP, CPE-G 侧的 VTEP 记为 G-VTEP。VxLAN 中的 VTEP 存有两个可达信息 MAP 表, 以确定上行封装过程中的目的地址, 分别是 VTEP L2 table 和 VLAN to VNID map。VLAN to VNID map 将 underlay 网络中的 VLAN 号映射至 overlay 网络的 VNID。而 VTEP L2 Table 是将 MAC 地址、VNID 和所在的 VTEP 一一对应。因此, 在上行链路 VTEP 封装时, 通过 VLAN to VNID map 表, 从 VLAN ID 查询到对方 VNID, 然后根据对方 MAC 地址和 VNID 得到对方所在的 VTEP 的 IP 地址; 而源端, 以自身所在的 VTEP 的 IP 地址为源地址, 而不是用户自己的 IP 地址。通过这样的方式, 可以完成隧道的封装。VTEP 解封时, 其过程与之相反。

隧道实现过程如下。

首先, 在 CPE-B 和 CPE-G 中部署 VTEP 后, B-VTEP 把用户的 Ethernet 数据帧 (frame) 封装到 UDP 报文中, 并送

到 IP 网络, 把这个过程称为上行链路的封装。然后, 在 IP 网络中, 经过封装的数据报文通过路由从 B-VTEP 传递到另一个 VTEP, 最后送到 G-VTEP。最后在 G-VTEP 中进行解封的操作, 并将符合以太网格式的用户数据送到用户。把这个过程叫做下行链路的解封操作。在这个过程中, 利用 VNID (VxLAN network identifier) 区分不同的用户和业务。可以通过 VNID 实现用户/业务的精细管理粒度。VNID 有 24 bit 的编码空间, 约为 16 MB, 大大超过 VLAN 的 12 bit 的编码空间^[6]。

3.1.2 VxLAN 技术隧道端点的实现方案

可以看到, 在采用 VxLAN 隧道技术后, 需要在原有 underlay 网络增加 VTEP, 而中间的其他网络可以保持不变。此外, 也可以增加 SDN 控制器, 利用 SDN 控制器管理 VTEP 间的路由和流量控制, 由于篇幅有限, 本文不讨论此种方案。由于接入网络需要满足带宽需求和业务需求各异的用户, 因此接入网的形态呈现多样化, 很难用统一的网络满足所有用户的需求。于是 VTEP 的部署位置也有多种选择, 以下是部署 VTEP 的几种方案^[7]。

B-VTEP 部署的方案有以下 3 种。

(1) 方案 1

将 B-VTEP 部署在 CPE 侧, 保留 CPE 部分三层功能, 更新现有 CPE 功能, 增加 VxLAN 隧道功能。

(2) 方案 2

将 B-VTEP 部署在 OLT 侧。CPE-B 仅保留二层功能, 在 OLT 增加 VxLAN 隧道功能。

(3) 方案 3

将 B-VTEP 部署在 BRAS (broadband remote access server, 宽带远程接入服务器) 侧。CPE 仅保留二层功能, 在 BRAS 增加 VxLAN 隧道功能。



G-VTEP 部署的方案有以下两种。

(1) 方案 1

将 G-VTEP 部署在 BRAS 侧。

(2) 方案 2

将 G-VTEP 部署在 CPE-G 侧的 CPE-MUX (CPE multiplexer) 内。CPE-MUX 是指 CPE-G 的入口点, 负责将 CPE-B 映射至 CPE-G^[4]。

考虑到简化 CPE 和利旧现有接入网络中处于路由模式和桥接模式的 CPE, 本文认为将 B-VTEP 放置于 OLT 侧或 BRAS 侧是更好的方案, 相应地, 可以根据实际情况将 G-VTEP 部署于 BRAS 侧或 CPE-G 侧的 CPE-Mux。也就是说 B-VTEP 更适合采用方案 2 和方案 3。原因如下。

首先, 如果 B-VTEP 采用方案 1, 则需要现有处于路由模式下的 CPE 增加 VxLAN 隧道功能, 且需要保留其部分三层功能。其次需要对现有的海量 CPE 进行升级改造或者替换, 因此存在成本高和可行性差的问题。最后桥接模式下的 CPE 不具备三层功能, 无法建立 VxLAN 隧道, 因此无法实现利旧。

总的来说, 虽然 B-VTEP 采用方案 1 可以保持现有接入网设备不变, 但存在以上的缺点, 而采用方案 2 和方案 3 尽管需要改造接入网以增加 VxLAN 功能, 但可行性更高, 实现难度更小。因此本文认为 B-VTEP 更适合采用方案 2 和方案 3。

虽然目前现有的 CPE、OLT 和 BRAS 等还不支持 VxLAN 协议, 近期看来用 VxLAN 技术实现 CPE-B 至 CPE-G 的连接的方案成熟度还不够, 但随着 VxLAN 应用越来越广泛并已有跨 IDC (internet data center) 实现二层互联互通的成功经验, 长远看来, 用 VxLAN 技术建立 CPE-B 与 CPE-G 的二层连接是未来趋势。

3.2 NvGRE 技术在 vCPE 中的应用

NvGRE 是一种 L2 over L3 的虚拟网络的隧道封装协议, 并且也提供了 24 bit 的网络标识号, 不受限于 VLAN 的 4 096 (12 bit) 的限制^[6]。基于 GRE (通用封装协议) 的 NvGRE 具有良好的兼容性, 但也有弊端。为了提供描述带宽利用率粒度的流, 传输网络需要使用 GRE 头, 但是这导致 NvGRE 不能兼容传统负载均衡, 这是 NvGRE 与 VxLAN 相比最大的区别也是最大的不足。与此同时, 在实现 CPE-B 与 CPE-G 间隧道端点间的路由和流量控制时, VxLAN 可以采用基于多播和泛洪的分布式方式, 也可以采用集中控制器的方法, 而 NvGRE 必须采用类似于 SDN

的集中控制器的方法。尽管采用集中控制器的方法具备更好的拓展性, 但这需要专用的软件和硬件支持, 这也是 NvGRE 面临的挑战之一。同样, vCPE 中采用 NvGRE 建立隧道时, 对于隧道端点的选择也有多种方式, 类似于上述 VxLAN 中隧道端点的选择, 本文不赘述。NvGRE 和 VxLAN 均可实现 CPE-B 和 CPE-G 的二层连接, 具体采用哪一种技术取决于各个厂商的利益考虑。以目前的情况来看, VxLAN 应用更为广泛, 应用前景更好。

4 vCPE 研究进展

目前业内正在积极开展 vCPE 研究。在标准化组织方面, 最为积极的是 BBF 组织, 于 2016 年 6 月份发布了 vCPE 技术标准^[4], 主要规范了 vCPE 技术要求和设计要求等方面。需要注意的是, 业内大多处于解决方案探索和设计阶段, 进展较为缓慢, 需要得到解决和统一规范的问题还很多, 因此 vCPE 离现网部署还有一定的距离。除了上文所述的 CPE-B 与 CPE-G 的连接问题, 还有以下问题需要进一步探讨。

(1) 安全性问题

由于 CPE 不再由家庭侧或企业侧一个单独的专用设备实现, 而是将其大部分功能部署于云端。而云端的资源是按需共享的, 不属于某个用户专用, 因此遭受 DoS、ALG 的风险也将提高。在部署 vCPE 时, 必须解决这些问题。

(2) 实时业务 QoS 保证的问题

CPE 网络虚拟化后, 由于将原家庭 LAN 侧的网络协议、配置管理和控制机制上收到网络侧, 一些对实时性要求较高的业务 (如语音、视频等) 的控制信令需要从本地发送到网络侧的 vCPE 进行处理, 与在本地 CPE 处理的方式相比, 此类实时性要求较高的业务体验有可能受到上下行带宽拥塞的影响, 因此网络虚拟化后需要为此类业务提供专门的 QoS 保障机制。

(3) 隐私问题

引入 vCPE 会使得家庭网络暴露给网络侧, 网络侧甚至可以根据需要对用户流量进行镜像和分析。家庭设备不再像传统方式隐藏于驻地侧 NAT 功能下, 因此会带来一些隐私性的问题。

(4) 管理维护问题

原有 CPE 是通过 ITMS+ 进行相应的远程控制、配置管理、故障管理、性能管理等。引入网络虚拟化后, 会对原有的管理方式造成影响, 因此需要对 TR-069 协议进行相应

的调整。

(5) 认证问题

原有的 CPE 设备必须通过 OLT (optical line terminal, 光线路终端) 和 BRAS 的 PPPoE 认证才能上网, 引入网络虚拟化后, 可能需要定义新的业务开通流程, 与传统认证方式的兼容也是个难题。

(6) 交互协议问题

虽然 CPE-B 仅保留较简单的二层协议功能, 但考虑到网络侧有远程配置、故障管理等实际需要以及部分业务的交互需要, 所以需要 CPE-G 到 CPE-B 之间的交互协议, 具体采用哪种方式与 vCPE 部署位置有一定的关系, 还需要进一步研究^[9]。

5 结束语

将 CPE 的三层功能上移至网络侧可简化 CPE。这种瘦终端加云网络的架构可以降低终端投入成本, 同时支持即插即用, 简化用户的操作。同时对运营商来说, 不仅可以管理家庭内网, 减少上门服务率, 降低运维成本, 还能为用户提供多样化的增值服务。本文主要阐述了 CPE-B 至 CPE-G 的连接方法, 通过选择合适的隧道端点, 在利旧现有桥接模式和路由模式的 CPE 的同时, 以对现有接入网络最小的改动的代价来实现 CPE-B 至 CPE-G 的连接。本文还总结了 vCPE 研究进展及所遇到的问题。由于 vCPE 研究领域较新颖, 业内还未开发出相应支持 overlay 技术的 CPE 和 OLT 等设备, 验证 CPE-B 至 CPE-G 的连接方法的性能较为困难。因此本文仅从理论上进行了分析, 未来的研究中将进行更多的实验验证。

参考文献:

- [1] 郑毅, 华一强, 何晓峰. SDN 的特征、发展现状及趋势[J]. 电信科学, 2013, 29(7): 84-88.
ZHENG Y, HUA Y Q, HE X F. Characteristics, development and future of SDN [J]. Telecommunications Science, 2013, 29(7): 84-88.
- [2] ETSI. Network functions virtualization(NFV): management and orchestration: ETSI. GS NFV-MAN 001 V1.1.1[S]. 2014.
- [3] BBF. Network enhanced residential gateway issue 1: BBF. TR-

317[S]. 2016.

- [4] 中国电信集团公司. 中国电信智能家庭网关技术规范: QCT2603-2015 [S]. 2015.
China Telecom Corporation. Technical specification for smart home gateway of China Telecom: QCT2603-2015[S]. 2015.
- [5] 沈成彬, 王成巍, 蒋铭, 等. 下一代 PON 技术的进展与应用[J]. 电信科学, 2010, 26(8): 1-7.
SHEN C B, WANG C W, JIANG M, et al. Progress & application of next generation PON technology [J]. Telecommunications Science, 2010, 26(8): 1-7.
- [6] WEERASINGHE J, ABEL F. On the cost of tunnel endpoint processing in overlay virtual networks [C]//IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC), December 8-11, 2014, London, UK. New Jersey: IEEE Press, 2014: 756-761.
- [7] MORAVCIK M, SEGEC P, PALUCH P, et al. Clouds in educational process [C]//2015 13th International Conference on Emerging eLearning Technologies and Applications (ICETA), November 26-27, 2015, Kezmarok, Slovakia. New Jersey: IEEE Press, 2015: 1-7.
- [8] 王瑾, 张沛. 公众客户网关虚拟化对网络运维的影响 [J]. 邮电设计技术, 2015(2): 24-27.
WANG J, ZHANG P. The influence of virtualized customer premise equipment to the network [J]. Designing Techniques of Posts and Telecommunications, 2015(2): 24-27.

[作者简介]



扶奉超(1991-),女,中国电信股份有限公司广州研究院数据通信工程师,主要从事宽带接入网技术及其虚拟化的研究工作。



谢元宝(1989-),男,中国电信股份有限公司广州研究院移动通信工程师,主要从事LTE网络关键技术的研究工作。