doi:10.3969/j.issn.1000-436x.2013.z2.029

动态 DNS+网关分离的分布式数据中心互联方法

谢胜军，蔡利平，殷锋
（西南民族大学 校园网络管理中心， 四川 成都 610041）

摘 要：分析了高校多校区分布式数据中心现存互联问题，在传统单侧网关和 RHI+网关分离互联方式的基础上，提出了一种 DDNS+网关分离的互联思路，以前端 DDNS 技术实现业务流量站点的选择，利用 SLB 设备配合 VMware 实现用户访问的不间断。在实验环境中，在线或新上线用户在虚拟机漂移前后能够快速访问到数据中心的同一种业务，适用于业务流量大、IP 规划复杂、连续性要求高的应用场景。该互联方法必须在数据中心的边缘增加 DDNS 设备，在两地的数据中心交换机侧旁挂 SLB 设备。

关键词：动态 DNS；分布式数据中心；网关分离
中图分类号：TP393 **文献标识码**：A **文章编号**：1000-436X(2013)Z2-0153-04

DDNS + gateway separation method for interconnect of distributed data center

XIE Sheng-jun, CAI Li-ping, YIN Feng
(Campus Network Management Center, Southwest University for Nationalities, Chengdu 610041, China)

Abstract: The multi-campus university distributed data centers interconnected problems existing in the traditional gateway were analyzed and based on RHI + gateway unilateral separation methods on the Internet, a DDNS + gateway interconnection separation idea to front-end technology was proposed to achieve business traffic sites DDNS selection, using the device with the VMware SLB to achieve uninterrupted user access. In a lab environment, online or new on-line users drifting around in a virtual machine can fast access to the data center to the same kind of business, suitable for the business flow, IP planning complex continuous demanding scenarios. The interconnection method must increase DDNS equipments in the data center edge, hanging the SLB device next to the switch side of data centers in both places.

Key words: DDNS; distributed data centers; gateway separation method

1 引言

国内高校多校区已经成为普遍现象，为了必须保证校内各种网络业务的一致性及漫游，分布式数据中心的建设也非常热门，2 个数据中心间的三层网络互联方式主要有传统的单侧网关模式[1]和路由健康注入（RHI）机制+网关分离模式[2]。

传统单侧网关模式是在 2 个数据中心间部署 VRRP[3]，调高主要数据中心的 VRRP 优先级，正常情况下通过主数据中心的网关设备转发各种业务流量，在出现故障后自动切换到备用的备用数据中心的网关设备。此模式下业务流量始终是从其中一个数据中心单进单出，流量路径明确，安全性强，但是数据中心之间受互联链路影响大，存在 VRRP 异常切换风险。

RHI+网关分离模式适用于业务流量大、连续性要求较高的场景，允许同一个网关同时部署在 2 个数据中心，业务流量在不过同情况下可分别从 2 个数据中心双进双出，避免了传统单侧网关模式存在的前段网络利用不高和依赖互联链路质量问题。此模式下几乎每台虚拟机都需要单独的主机路由，如果数据中心虚拟机较多，就会导致网络的路由条目庞大，严重影响整个数据中心的稳定性。

为了解决以上 2 种互联模式存在的问题，本文

收稿日期：2013-09-05
基金项目：国家自然科学基金资助项目(61379019)
Foundation Item: The National Natural Science Foundation of China(61379019)

作者提出了一种有别于以上 2 种数据中心的互联方式，利用动态 DNS[4]和网关分离技术相结合来实现分布式数据中心的互联。

2 互联方法设计

该互联方法适用场景与“RHI+网关分离”类似，不同的是采用该方法互联的数据中心侧必须配置动态 DNS 设备或系统，通过动态 DNS 可以实现访问流量的智能分发与调配，无需增加广域网开销。该方法涉及到 2 个关键技术，一是“网关分离”，就是 2 个数据中心部署同一个网段的网关，网关的 IP 地址相同，虚拟机就近选择所属数据中心的网关来完成业务流量的三层转发；另一关键技术是“动态 DNS”（DDNS）技术[4]，承载同一个业务的虚拟机在不同数据中心通过 NAT 由负载均衡设备（SLB）实现呈现不同的服务 IP 地址。如果虚拟机进行了迁移即可触发 vCenter 上的可执行脚本，自动修改 DDNS 上该虚拟机 IP 地址对应的 DNS 记录，由此实现新上线用户的访问业务流量的三层路径的优化。

2.1 互联方法的拓扑

采用本方法进行互联的数据中心拓扑结构与其他 2 种互联模式有所不同，需要在其中一个数据中心的网络边界增加 DDNS 设备。如图 1 所示，数据中心 1（IDC1）和数据中心 2（IDC2）的服务器采用了 VMware 虚拟化技术[5]，vCenter 部署在 IDC1，数据中心的汇聚层实现二层互联，2 个数据中心同时部署同一个网段的网关，网关 IP 地址相同，虚拟机就近选择本数据中心的网关进行三层流量转发；汇聚层设备上旁挂主备方式部署的 SLB 设备，汇聚于核心路由器间部署主备方式的防火墙；边界部署了动态 DNS 设备，用户端通过域名方式访问数据中心的业务系统。

当通过 IDC1 中的 vCenter 将业务系统 web1.edu.cn 对应的虚拟机从 IDC1 迁移至 IDC2 时，当前已在线用户的访问流量不中断，而新上线的用户则选择三层最优路径访问位于 IDC2 中的 web1.edu.cn 业务。

2.2 在线用户在虚拟机漂移前的流量路径

如图 1 所示，当 web1.edu.cn 所对应的虚拟机还没有进行迁移动作，仍处于 IDC1 中时，用户通过终端访问 web1.edu.cn 是的流量路径如下。

Step1 用户需要访问 web1.edu.cn 时，其终端向网络边界部署的 DDNS 服务器发起域名查询请求，最终返回查询结果显示 web1.edu.cn 对应的 IP 地址是 IDC1 中 SLB1 上配置的 VIP-1。

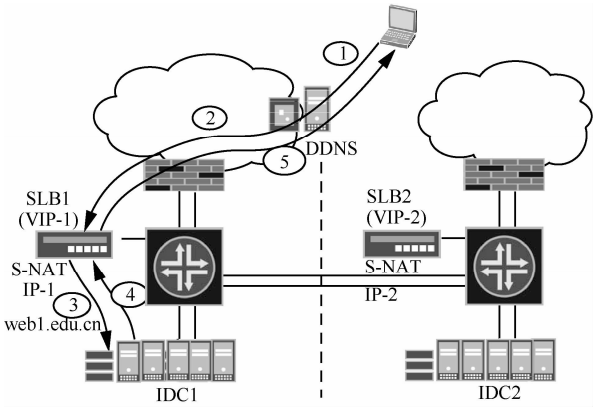


图 1 动态 DNS+网关分离

Step2 用户终端发起的流量经过核心设备的转发来到 IDC1 的 SLB1 主设备上。

Step3 SLB1 设备对用户终端的流量做 NAT（地址转换），报文的源 IP 被改为 SLB1 的接口地址 IP-1 地址，目的地址被改为虚拟机的真实地址 VM-IP。

Step4 虚拟机到用户终端回程报文的源 IP 是 VM-IP，目的 IP 是 SLB1 的接口地址 IP-1。由于虚拟机的网关指向汇聚交换机，从虚拟机到用户终端的回程流量经汇聚交换机转发到主 SLB1 上。

Step5 SLB1 查询会话表后，将该 IP 报文的源地址改为 VIP-1，将目的地址改为用户终端的实际 IP，最后该 IP 报文经核心设备转发至用户终端。

2.3 在线用户在虚拟机漂移后的流量路径

如图 2 所示，如果因为某种情况需要将 IDC1 中的 web1.edu.cn 对应的虚拟机通过 vCenter 迁移至 IDC2 中，此时由于用户的终端设备仍然具有本地的 DNS 缓存，在 DNS 缓存超时之前，用户终端仍然将 web1.edu.cn 解析成 VIP-1。由于从用户终端到提供 web1.edu.cn 业务的虚拟机的访问路径上存在防火墙，且防火墙通常采用基于 TCP/UDP 状态的会话检查机制，所以对于已经在线的用户（防火墙上已建立用户到虚拟机的会话表项）必须保证虚拟机迁移后，从用户到虚拟机的访问仍然保持原有路径。这就意味着从用户访问 web1.edu.cn 的数据流量在 Step2 和 Step3 会发生变化，分别如 Step2′和 Step3′所示。

Step2′ SLB1 上经过 NAT 处理后的报文被转发至 IDC 1 的汇聚交换机（有 VM-1 对应的直连网段路由），该汇聚交换机查询 ARP 表后，将报文发

往与 IDC2 的汇聚交换机相连的端口，IDC2 的汇聚交换机查询 MAC 表，完成最终转发。

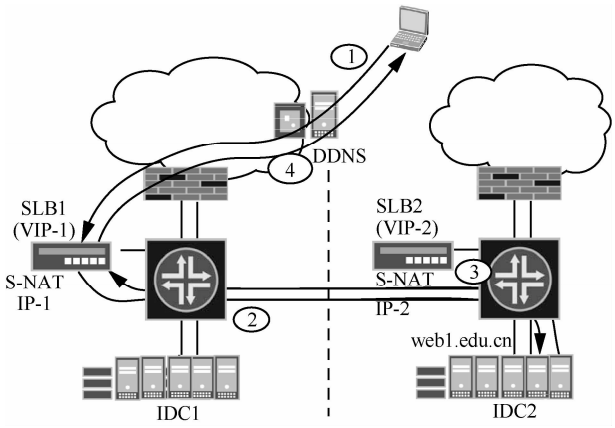


图 2　在线用户的虚机迁移后流量

Step3′　由于 IDC1 的汇聚设备与 IDC2 的汇聚设备采用了网关分离部署方式，所以虚拟机到用户终端的回程报文首先发向 IDC2 的本地网关（IDC2 的汇聚交换机）。IDC2 汇聚交换机与 IDC1 汇聚交换机互联的三层接口上已经学到 SLB1 接口地址 IP-1 对应的路由，所以回程报文经 IDC1 汇聚交换机转发后回到 SLB1。

这种次优路径流量不会一直存在，当用户结束对 web1.edu.cn 的所有访问流量后一段时间，本地 DNS 缓存超时清空，用户如果再次发起 TCP/UDP 会话，则业务流量将按照 2.4 节中所述的新上线用户的流量路径完成转发。

2.4　新上线用户在虚拟机漂移后的流量路径

VMware vCenter 支持基于事件触发的脚本技术，管理员可以针对 vCenter 上发生的多种类型的事件（Event）定义执行脚本（TCL/TK）。本互联方法中，必须在 vCenter 中针对 web1.edu.cn 对应虚拟机从 IDC1 到 IDC2 的动态迁移事件定义一个执行脚本，核心内容就是“Telnet 到 DDNS 设备上，将 web1.edu.cn 的 DNS 解析改为 VIP-2”。所以如图 3 所示，通过 vCenter 将 web1.edu.cn 对应的虚拟机从 IDC1 中迁移至 IDC2 后，新上线用户的流量路径如下。

Step1　当 web1.edu.cn 对应虚拟机由 IDC1 迁移至 IDC2 后，通过 vCenter 触发 DDNS 上的域名解析变更脚本，将 DDNS 上 web1.edu.cn 域名的解析地址由 VIP-1 变更为 VIP-2。

Step2　新上线的用户访问 web1.edu.cn，其用户终端向 DDNS 服务器发起查询请求，最终返回查询结果，web1.edu.cn 对应的地址是 IDC2 中的 SLB2 上配置的 VIP-2。

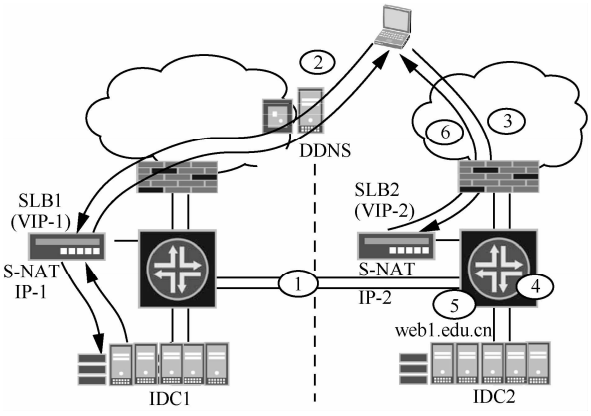


图 3　新上线用户的虚机迁移后流量

Step3~Step6 与 2.2 节中用户第一次访问 IDC1 中 web1.edu.cn 的 Step2~Step5 相同，只不过是通过 IDC2 中的 SLB2 等设备来完成相关业务流量的转发动作。

3　小结及对比

本互联方法是通过数据中心网络前端的 DDNS 设备实现业务流量站点的选择，在数据中心内部通过 NAT 可访问呈现不同 IP 地址的同一虚拟机，进而实现业务流量的智能调配，无需额外发布路由。虚拟机迁移后在客户端 DNS 缓存老化时间内存在次优路径，通过与 vCenter 的联动可保证迁移后新建连接以优化路径实现访问。

本互联方法与传统单侧网关、RHI+网关分离 2 种互联模式对比如表 1 所示。

以上 3 种互联方式各有优缺点，互相并不排斥。在规划校园网数据中心时，应按照实际业务需要，对于业务流量不大、便于区分、业务连续性要求不高的场景采用单侧网关方案；而对于业务流量较大、IP 规划复杂、业务连续性要求较高的场景采用后 2 种网关分离的方案。

4　结束语

传统单侧网关模式无论是从技术实现还是从配置维护的角度来说，都是最为简单的，但其无法适应现在大规模数据中心的复杂场景。“DHI+网关分离”模式所需条件虽然较为简单，但发布主机路由的方式会导致广域网开销的上升，且禁用防火墙状态监测，数据中心的安全无法得到保障。而“动态 DNS+网关分离”的互联方法则只

表 1　　**3 种分布式数据中心互联模式比较**

比较模型	适用场景	技术成熟度	投资成本	管理难易度	业务健壮度	优点	缺点
单侧网关	业务流量小、便于区分、连续性要求不高	高	低	简单	低	流量路径明确、配置简单；安全性强，FW 基于状态监测报文	基于业务网段的不同来区分，粒度较大；当互联链路存在链路质量问题时，次优路径影响较为明显
RHI+网关分离	业务流量大、IP 规划复杂、连续性要求高	较高	较低	较难	低	网关分离，无需跨中心运行 VRRP，网络故障域只限定在单侧数据中心；通过主机路由注入，解决了"次优路径"问题，降低了互联链路的带宽占用；可根据某些业务或者 IP 地址进行流量的分担，控制更加灵活便利	防火墙须关闭状态监测功能，安全性低；路由检测设备成为关键节点，可靠性和可维护性低；有些用户由于安全监管的要求，不允许增加检测设备，或者不允许发布主机路由，导致方案无法部署；引入大量主机路由，会对整网路由的稳定性和自愈时间造成冲击，网络维护效率较低
动态 DNS+网关分离	业务流量大、IP 规划复杂、连续性要求高	较高	较低	简单	高	用户基于 DNS 访问对应的业务系统；依靠 SLB 设备的 NAT 功能，同一个虚拟机在不同的数据中心呈现不同的 VIP；SLB 对来自客户端的流量做 S-NAT，实现网络会话的连续不中断；vCenter 上通过虚拟机迁移事件触发执行脚本以修改 DNS 对域名的解析	NAT 后源地址转换，服务器无法获知客户端实际 IP；数据中心侧必须有 DDNS 设备

需要在数据中心侧配置 DDNS 设备，无需增加广域网开销，以少许的成本投入换取了业务的顽健和后期维护的简单高效。总之，在分布式数据中心的设计规划中，应根据高校自身的业务需求、投资额度、实现难度等因素，灵活构建更符合实际使用环境的网络互联方式。

参考文献：

[1] 黄志强.数据中心统一建构网络设计构想[J]. 信息技术, 2012(11): 58-59.
HUANG Z Q. The data center and unified design idea of the network[J]. Journal of Information Technology, 2012 (11) : 58-59.

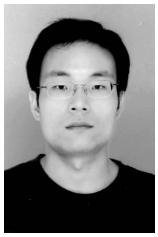
[2] 李阳阳, 王洪波, 张鹏等. 基于多属性信息的数据中心间数据传输调度方法[J]. 通信学报, 2012, 33(Z1):121-131.
LI Y Y, WANG H B, ZHANG P, *et al*. Based on multiple attribute information data center between data transmission scheduling method[J]. Journal of Communications, 2012, 33(Z1): 121-131.

[3] RFC3768.Virtual Router Redun Dancy Protocol(VRRP)[R]. 2004.

[4] 鄢萍, 易润忠, 童亮. 基于DDNS和NAT的服务器内外网动态映射[J].计算机工程, 2008(20):136-137.
YAN P, YI R Z, TONG L. Based on DDNS and NAT server dynamic mapping between inner and outer net[J]. Computer Engineering, 2008, (20): 136-137.

[5] 何禹, 胡宇鸿, 王一波. 虚拟化技术在校园网数据中心的应用[J]. 电子科技大学学报, 2007, (36):1461-1464.
HE Y, HU Y H, WANG Y B. Virtualization technology application in campus network data center[J]. Journal of University of Electronic Science and Technology, 2007, (36):1461-1464.

作者简介：



谢胜军（1977-），男，四川成都人，硕士，西南民族大学校园网络管理中心工程师，主要研究方向为网络运行管理。

蔡利平（1973-），女，四川成都人，硕士，西南民族大学校园网络管理中心高级工程师，主要研究方向为计算机应用软件开发。

殷锋（1972-），男，四川成都人，博士，西南民族大学校园网络管理中心教授，主要研究方向为计算机应用软件测试。