计算机科学
COMPUTER SCIENCE

基于虚拟化的跨域VPN解决方案

陶志勇, 张锦, 阳王东

引用本文

陶志勇, 张锦, 阳王东. 基于虚拟化的跨域VPN解决方案[J]. 计算机科学, 2023, 50(9): 357-362.

TAO Zhiyong, ZHANG Jin, YANG Wangdong. Solution to Cross-domain VPN Based on Virtualization [J]. Computer Science, 2023, 50(9): 357-362.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.220800252

基于虚拟化的跨域 VPN 解决方案

陶志勇[1,2]　张　锦[2,3]　阳王东[2]
1 长沙民政职业技术学院软件学院　长沙 410004
2 湖南大学信息科学与工程学院　长沙 410082
3 湖南师范大学信息科学与工程学院　长沙 410012
(27537406@qq.com)

摘　要　针对目前运营商网络中构建的跨域虚拟私有网实现复杂、自治系统边界设备负载过重、存在单点故障等问题,提出了采用虚拟化方式构建跨域虚拟私有网的解决方案。该方案包括公网隧道的建立、本地 VPN 实例的建立、自治系统边界设备的虚拟化、边界设备私网路由的交互 4 个关键步骤。为评估方案的可行性,对方案进行了测试与验证,测试与验证结果表明该方案达到了预期设计的目标。为了评估方案的优越性,与传统多跳 EBGP 方式构建的跨域虚拟私有网在交换容量、路由条目、标签条目等维度进行了对比分析。对比结果表明,采用该方案构建的跨域虚拟私有网增强了自治系统边界设备的数据处理能力,并减少了自治系统边界设备需处理的数据量,是一种构建跨域私有网的改进方案。

关键词:虚拟化;多协议标签交换;边界网关路由协议;自治系统边界设备;虚拟私有网
中图法分类号　TP393

Solution to Cross-domain VPN Based on Virtualization

TAO Zhiyong[1,2],ZHANG Jin[2,3] and YANG Wangdong[2]
1 Software School,Changsha Social Work College,Changsha 410004,China
2 College of Computer Science and Electronic Engineering,Hunan University,Changsha 410082,China
3 College of Computer Science and Electronic Engineering,Hunan Normal University,Changsha 410012,China

Abstract　To address the problems of complex implementation of cross-domain virtual private networks built in current carrier networks,excessive load on devices at the border of autonomous systems,and the existence of single points of failure,this paper proposes a solution for building cross-domain virtual private networks by virtualization. The scheme consists of four fundamental steps:the establishment of public network tunnels,the establishment of local VPN instances,the virtualization of autonomous system border devices,and the interaction of private network routes of border devices. To evaluate the feasibility and superiority of the scheme,comparative experiments are conducted with the cross-domain virtual private network constructed by the traditional multi-hop EBGP approach in the dimensions of switching capacity,route entries,and label entries. Experimental results show that the cross-domain virtual private network constructed by this scheme enhances the data processing capability of the autonomous system boundary devices and reduces the amount of data to be processed by the autonomous system boundary devices. In general,this improved scheme is advanced and effective for building cross-domain virtual private networks.

Keywords　Virtualization,Multi-protocol label switching,Border gateway routing protocol,Autonomous system boundary sevice,Virtual private network

1　引言

因特网工程组(Internet Engineering Task Force,IETF)发布的 RFC-4364,阐释了在运营商的网络中采用多协议标签交换[1](Multiprotocol Label Switching,MPLS)与边界网关路由协议[2](Border Gateway Protocol,BGP)为用户构建虚拟私有网[3](Virtual Private Network,VPN)的解决方案。因该方案构建的虚拟私有网能有效隔离不同用户的私网数据,且稳定性与扩展性好,使得该方案构建的虚拟私有网的应用越来越广泛[4-6]。

到稿日期:2022-08-27　返修日期:2023-01-18
基金项目:国家自然科学基金(61872127);湖南省教育厅资助科究项目(22C1433);湖南省自然科学基金(2020JJ7089);长沙民政职业技术学院横向项目(HX2023025)
This work was supported by the National Natural Science Foundation of China(61872127),Research Foundation of the Education Department of Hunan Province(22C1433),Natural Science Foundation of Hunan Province,China(2020JJ7089)and Horizontal Project of Changsha Social Work College(HX2023025).
通信作者:张锦(mail_zhangjin@163.com)

随着用户业务的不断扩张，其用户网络需跨不同自治系统[7]（AutonomousSystem，AS）。而 RFC-4364 所提供的 MPLS 与 BGP 构建的虚拟私有网局限于在同一自治系统内（域内），不能跨自治系统（域间）为用户提供服务。因此，研究跨不同自治系统来构建虚拟私有网，是亟待解决的问题[8]。

为了构建跨域的虚拟私有网，目前主要的解决方式有 3 种：1）背靠背方式，该方式在自治系统边界设备上为每一个用户提供单独的接口来构建跨域私有网；2）单跳的 EBGP 方式，该方法在自治系统边界设备上传递扩展的 BGP 路由信息来构建跨域私有网；3）多跳的 EBGP 方式，该方式在连接用户的边界设备上传递扩展的 BGP 路由信息来构建跨域私有网[9-10]。上述 3 种方式都存在一定的缺陷，背靠背与单跳的 EBGP 方式使得自治系统边界设备不但需处理私网的数据，而且需处理用户的公网数据；而多跳的 EBGP 自治系统边界设备只需处理公网数据，私网数据由连接用户的边界设备来处理，减轻了自治系统边界设备的负担，但该方式需修改 MPLSVPN 体系框架，实现复杂。因此，上述 3 种方式有待进一步改进[11-12]。

在运营商的网络中，一台自治系统边界设备需为上万甚至十万用户提供 VPN 的服务。背靠背、单跳的 EBGP、多跳的 EBGP 方式都存在设备负载过重的问题，负载过重会影响用户业务数据的交互，严重时会导致网络瘫痪。为此，本文提出了一种虚拟化的跨域 VPN 解决方案[13-14]。

IRF[15]（Intelligent Resilient Framework）是一种网络设备虚拟化技术，以 IRF 智能弹性架构为关键字在知网上进行检索（截止检索日期为 2022 年 7 月 10 日），检索到学术期刊 21 篇，学位论文 4 篇，会议论文 1 篇[16]。以上数据表明，该技术的研究还处于起步阶段。而本文针对背靠背、单跳 EBGP、多跳 EBGP 构建的跨域 VPN 都存在局限性的问题，提出了一种虚拟化的跨域 VPN 解决方案。该方案将通过网络设备虚拟化技术将多台自治系统边界设备虚拟成一个资源池，让资源池中的多台自治系统边界设备共同分担公网数据与私网数据的处理与传送，并通过 MPLS 与 BGP 技术为不同用户构建其独立的虚拟私有网，在隔离不同用户数据的同时，实现其私网数据的交互。同时，将所提方案与传统方式构建的跨域 VPN 在包转发率、处理与维护路由条目数、处理与维护的标签条目数等维度进行对比，以验证该方案的优越性[17-18]。

2 基于虚拟化的跨域 VPN 的方案设计

2.1 设计理念

虚拟化的跨域 VPN 构建采用三层虚拟的设计理念，第一层虚拟是依托网络设备虚拟化技术将多台自治系统边界设备构建为一个资源池，由资源池中的设备负载分担公网数据与用户私网数据的识别与传输。第二层虚拟是借助 MPLS 在公网中建立虚拟私有网，为不同用户的私网数据穿越公网提供通道。第三层虚拟是利用 BGP 在边界设备与自治系统边界设备上给不同用户分配不同的私网标签，为不同用户构建其对应的虚拟网络。

2.2 网络模型

为了验证设计理念的可行性，所提方案构建了一个实验所需要的网络模型，如图 1 所示。图 1 所示的网络模型需在自治系统间实现 *A* 用户总部与分部、*B* 用户总部与分部的私网数据的交互。

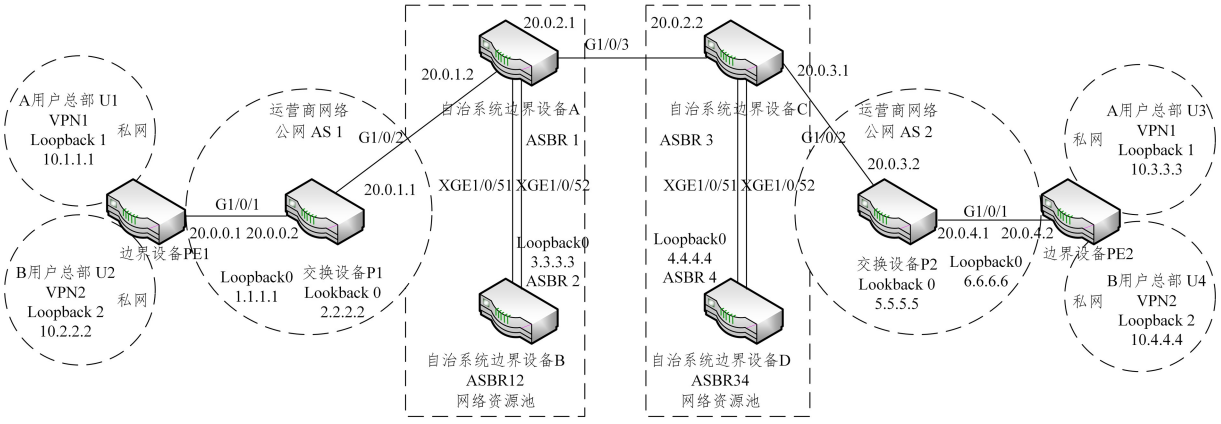

图 1 网络模型

Fig. 1 Network Model

为了后续阐述的方便，对网络模型中的设备进行定义。*A* 用户总部与 *B* 用户总部用 loopback 地址来模拟，分别用 U_1 与 U_2 表示，*A* 用户分部与 *B* 用户分部用 loopback 地址来模拟，分别用 U_3 与 U_4 表示；连接 *A* 用户总部与 *B* 用户总部的运营商边界设备用 PE_1 表示，连接 *A* 用户分部与 *B* 用户分部的运营商边界设备用 PE_2 表示；自治系统 1 的运营商交换设备表示为 P_1，自治系统 2 的运营商交换设备表示为 P_2；自治系统 1 的运营商自治系统边界设备 *A* 与 *B* 用 $ASBR_1$ 与 $ASBR_2$ 表示，自治系统 2 的运营商自治系统边界设备 *C* 与 *D* 用 $ASBR_3$ 与 $ASBR_4$ 表示；运营商自治系统边界设备 *A* 与 *B* 虚拟化后的设备表示为 $ASBR_{12}$，运营商自治系统边界设备 *C* 与 *D* 虚拟化后的设备表示为 $ASBR_{34}$；*A* 用户总部与分部构建的虚拟私有网表示为 VPN_1，*B* 用户总部与分部构建的虚拟私有网表示为 VPN_2。

2.3 方案设计

实现实验模型中两个用户总部与分部的私网数据交互，在自动系统内与自治系统间均需构建传输私网数据的传输通道。在自治系统内使用 MPLS 的标签分发协议（Lable Distribution Protocol，LDP）分配公网标签，形成承载私网数据的标签交换路径（Lable Switching Path，LSP）。采用扩展的

MBGP 协议给不同用户私网数据分配私网标签，通过该私网标签形成 VPN-LSP，识别与区分不同用户的私网数据。在自治系统间传输私网数据时，由虚拟化后的自治系统边界设备 $ASBR_{12}$ 与 $ASBR_{34}$ 负载分担数据，并在 $ASBR_{12}$ 与 $ASBR_{34}$ 上给每个用户创建传输私网数据的专属接口，并通过 BGP 来交互不同用户的私网数据，进而实现两个用户总部与分部的私网数据交互。两个用户在自治系统内与自治系统间的数据交换过程如图 2 所示。

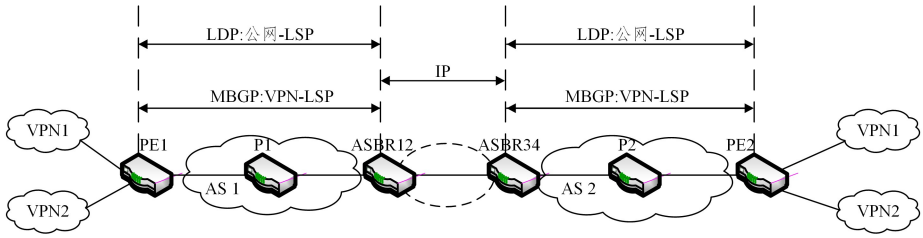

图 2　私网用户的数据交换过程

Fig. 2　Data exchange process of private network users

保障图 2 中两个用户总部与分部的私网数据交互，在实验模型中需完成公网隧道的建立、本地 VPN 实例的建立、自治系统边界设备的虚拟化、边界设备私网路由的交互 4 个关键步骤。

1）公网隧道的建立

公网的边界设备 PE_1 与 PE_2 能识别公网与私网数据，而位于公网中的 P_1 与 P_2 无私网的路由信息，故识别不了 A 用户与 B 用户的私网数据。因此，需在自治系统内的 PE_1，P_1，$ASBR_{12}$ 与 PE_2，P_2，$ASBR_{34}$ 设备上部署 MPLS，分配标签，形成标签转发表，并将私网数据封装在 MPLS 的标签内，使 P_1 与 P_2 读取外层 MPLS 标签来转发数据，进而通过 MPLS 生成的标签在公网中建立一条逻辑通道来承载不同用户的私网数据。公网隧道的建立需在设备上完成 3 个方面的部署，以自治系统 1 为例。首先启用 PE_1，P_1，$ASBR_{12}$ 设备的 MPLS 功能，其次在 PE_1，P_1，$ASBR_{12}$ 上运行 mplslsr-id，使用设备的 loopback 接口来标识设备的身份，再次在 3 台设备的接口上启动 MPLS 与 MPLSLDP 协议，给 loopback 接口分配标签，并使接口具备标签转发的能力。自治系统 2 的公网隧道的建立的部署与自治系统 1 相同，在此不再赘述。

2）本地 VPN 实例的建立

公网隧道的建立为 A 用户与 B 用户的私网数据穿越公网夯实了基础，而两个用户私网数据的识别与隔离需在 PE_1 与 PE_2、$ASBR_{12}$ 与 $ASBR_{34}$ 上建立本地的 VPN 实例。在设备上运行 ipvpn-instance 指令给两个用户建立本地的 VPN 实例，并在 VPN 实例下执行 route-distinguisher 与 vpn-target 指令，为两个用户的私网数据打上不同的标记，并通过多进程与虚拟路由技术为每一个用户在 PE_1 与 PE_2、$ASBR_{12}$ 与 $ASBR_{34}$ 上建立独立的实例路由表，进而分辨出不同用户私网数据，并实现不同用户私网数据的相互隔离。

3）自治系统边界设备的虚拟化

位于自治系统边界的设备不但需承载公网数据，而且需承载私网数据，当用户数呈上万增长时，负载过重，影响业务数据的正常交互。故需要网络设备虚拟化技术 IRF 将自治系统边界 $ASBR_1$ 与 $ASBR_2$、$ASBR_3$ 与 $ASBR_4$ 虚拟化，虚拟化后由两台自治系统边界设备共同负载分担数据的接收与传送。而实现 $ASBR_1$ 与 $ASBR_2$、$ASBR_3$ 与 $ASBR_4$ 来负载分担用户数据，首先需通过 irf member 指令分别给 $ASBR_1$ 与 $ASBR_2$、$ASBR_3$ 与 $ASBR_4$ 赋予不同的成员编号，标识其设备在资源池中的身份，同时方便后续通过成员编号来管理设备；其次使用 irf-port 指令将图 1 中 $ASBR_1$ 与 $ASBR_2$、$ASBR_3$ 与 $ASBR_4$ 相连的物理端口加入到虚拟化的逻辑端口下，使用逻辑端口下的物理端口交互虚拟化所需的协议报文；最后在 $ASBR_1$ 与 $ASBR_3$ 上执行 irf member 1 priority 10 指令，操控 $ASBR_1$ 与 $ASBR_3$ 成为资源池中的 Master 设备。通过上述步骤分别在 $ASBR_1$ 与 $ASBR_2$、$ASBR_3$ 与 $ASBR_4$ 上构建了一个资源池，可以在 Master 设备上通过 irf switch-to member-id 指令管理与调度资源池中从设备的资源，并由两台设备共同承担公网数据与私网数据的传输，实现数据的负载分担。后续如资源池中的两台设备随着用户数倍增，则还存在数据负载过重的问题，可以在资源池中再添加设备，新增设备到资源池时，支持“热插拔”，不影响资源池的正常运行，无须中断业务。

4）边界设备私网路由的交互

实现图 1 中 U_1 与 U_3、U_2 与 U_4 的私网数据的交互分为两个部分：在域内和在域间。考虑到实验模型中 P_1 与 P_2 不能学习私网数据的路由信息，故采用 BGP 路由协议。因 BGP 路由协议可以跨设备建立邻居关系交互私网数据的路由信息，所以在自治系统内的 PE_1 与 $ASBR_{12}$、PE_2 与 $ASBR_{34}$ 上采用扩展后的 MP-BGP 交互私网数据的路由信息，而在自治系统间的 $ASBR_{12}$ 与 $ASBR_{34}$ 上也采用 BGP 来交互私网数据的路由信息。首先在自治系统 1 的边界设备 PE_1 与 $ASBR_{12}$ 上运行 BGP 协议后，执行 peer 与 address-family ipv4 unicast 指令建立普通的 BGP 邻居关系，然后执行 address-family vpnv4 建立扩展的 BGP 邻居关系，最后执行 ip vpn-instance vpn1 与 ip vpn-instance vpn2，并通过 import-route direct 指令将 U_1 与 U_2 的私网数据分别引入到 vpn1 与 vpn2 的实例中，实现将 U_1 与 U_2 的私网数据由 PE_1 传送给 $ASBR_{12}$，在自治系统 1 内完成了 U_1 与 U_2 私网数据的路由交互。自治系统 2 内 U_3 与 U_4 的私网数据的路由信息由 PE_2 传送给 $ASBR_{34}$ 的方式与上述部署相同，在此不再赘述。

上述部署完成了不同用户私网数据在域内边界设备的交互，在域间的私网数据路由信息的交互在 $ASBR_{12}$ 与 $ASBR_{34}$ 上给每个用户创建专属的接口，并在该接口下通过 BGP 协议完成不同用户私网数据路由信息的交互。具体的部署是先在 $ASBR_{12}$ 与 $ASBR_{34}$ 相连接的主接口下创建两个子接口，并在子接口下执行 ip binding vpn-instance 指令让两个子接口分别与实例 vpn1 与 vpn2 绑定，然后在 BGP 路由协议的 ip vpn-

instance vpn1 与 vpn2 下，通过 peer 指令在其用户相对应的专属子接口下建立 BGP 邻居关系，并在专属的子接口下交互两个用户的私网数据的路由信息，进而实现 U_1 与 U_3、U_2 与 U_4 的私网数据的交互。

当运营商网络承载用户数不多时，采用该方式来部署跨域 VPN 的实现过程简单，管理方便，不需要修改协议与 MPLS 的体系结构。但该方式存在 3 个缺点：(1)扩展性差，有多少个用户则需在 $ASBR_{12}$ 与 $ASBR_{34}$ 上建立多少个本地的 VPN 实例，随着用户的增多，工作量成倍增长；(2)有多少个用户就需在 $ASBR_{12}$ 与 $ASBR_{34}$ 相连的主接口下创建多少个用户的专属接口与 IP 地址，接口与地址占用率高；(3)$ASBR_{12}$ 与 $ASBR_{34}$ 相连的主接口若不能创建逻辑接口，则该方案无法实施。

针对上述问题，对虚拟化的跨域 VPN 解决方案进行了改进。改后的方案公网隧道的建立、自治系统边界设备的虚拟化与改前的解决方案相同。不同之处在于，在 $ASBR_{12}$ 与 $ASBR_{34}$ 上无须创建交互私网数据路由信息的专属接口，也不需要在 $ASBR_{12}$ 与 $ASBR_{34}$ 建立本地的 VPN 实例，只需在 $ASBR_{12}$ 与 $ASBR_{34}$ 上建立 MP-BGP 的邻居关系，并将用户私网数据的路由信息在 VPNV4 下进行交互，且启动 $ASBR_{12}$ 与 $ASBR_{34}$ 相连接口的标签转发功能，进而完成不同用户私网路由信息在 $ASBR_{12}$ 与 $ASBR_{34}$ 间的交互。具体的部署是进入 $ASBR_{12}$ 与 $ASBR_{34}$ 相连的接口，执行 mpls enable，使接口具有处理标签的能力，然后在 BGP 路由协议视图下执行 address-family vpnv4，并通过 peer 指令建立 VPNV4 的邻居关系，最后通过 undo policy vpn-target 指令将域内的私网数据路由信息通过 BGP 路由协议在域间的 $ASBR_{12}$ 与 $ASBR_{34}$ 上进行交互，进而完成 U_1 与 U_3、U_2 与 U_4 的私网数据的交互。

3 性能评估

3.1 方案验证

为验证方案部署是否达到了预期的目标，对方案进行相应的验证。

1)网络设备虚拟化的建立

在 $ASBR_1$ 与 $ASBR_3$ 上通过执行 disirf 指令查看 $ASBR_1$ 与 $ASBR_2$、$ASBR_3$ 与 $ASBR_4$ 是否虚拟化成功，图 3 所示的 master 与 standby 的状态充分说明自治系统边界设备 $ASBR_1$ 与 $ASBR_2$、$ASBR_3$ 与 $ASBR_4$ 已成功虚拟化，两台设备能共同承担公网数据与私网数据的处理与传输。

```
<ASBR1>dis irf
MemberID    Role     Priority  CPU-Mac         Description
 *+1        Master   1         2c08-b2ae-0204  ---
   2        Standby  1         2c08-ad6d-0104  ---
--------------------------------------------------
 * indicates the device is the master.
 + indicates the device through which the user logs in.

 The bridge MAC of the IRF is: 2c08-b2ae-0200
 Auto upgrade                : yes
 Mac persistent              : 6 min
 Domain ID                   : 100
<ASBR1>
<ASBR3>dis irf
MemberID    Role     Priority  CPU-Mac         Description
 *+1        Master   1         2c08-c6fd-0504  ---
   2        Standby  1         2c08-ca27-0604  ---
```

图 3 虚拟化状态结果图

Fig. 3 Virtualization status results graph

2)公网隧道的 LSP 的建立

在 $ASBR_1$ 与 $ASBR_3$ 上执行 dis mpls ldp lsp 指令，得到图 4 所示的状态结果，图 4 所示的状态结果显示 PE_1，P_1，$ASBR_1$，PE_2，P_2，$ASBR_3$ 的 loopback 接口 MPLS 已分配了公网标签。至此，图 2 所示的自治系统内的公网 LSP 已形成，为 U_1 与 U_3、U_2 与 U_4 的私网数据穿越公网提供了通道。

```
<ASBR1>dis mpls ldp lsp
Status Flags: * - stale, L - liberal, B - backup
FECs: 3             Ingress: 2          Transit: 2          Egress: 1

FEC                 In/Out Label        Nexthop         OutInterface
1.1.1.1/32          -/1151              20.0.1.1        GE1/0/2
                    1151/1151           20.0.1.1        GE1/0/2
2.2.2.2/32          -/3                 20.0.1.1        GE1/0/2
                    1150/3              20.0.1.1        GE1/0/2
3.3.3.3/32          3/-
                    -/1150(L)

<ASBR1>
<ASBR3>dis mpls ldp lsp
Status Flags: * - stale, L - liberal, B - backup
FECs: 3             Ingress: 2          Transit: 2          Egress: 1

FEC                 In/Out Label        Nexthop         OutInterface
4.4.4.4/32          3/-
                    -/1150(L)
5.5.5.5/32          -/3                 20.0.3.2        GE1/0/2
                    1151/3              20.0.3.2        GE1/0/2
6.6.6.6/32          -/1151              20.0.3.2        GE1/0/2
                    1150/1151           20.0.3.2        GE1/0/2
```

图 4 公网的 LSP 状态结果

Fig. 4 LSP status results of public network

3)域内与域间私网的 VPN-LSP 路径的建立

私网的 VPN-LSP 用来识别与区分不同用户的私网数据，其 inlable 是设备分配给别的设备使用的标签，而 outlable 是别的设备分配给本设备使用的标签。在 $ASBR_1$ 与 $ASBR_3$ 上执行 dis bgp routing-table vpnv4 inlable | begin Network 后，其私网的 VPN-LSP 状态结果如图 5 所示。

```
<ASBR1>dis bgp routing-table vpnv4 inlabel | begin Network
     Network          NextHop         OutLabel        InLabel

*  >i 10.1.1.1/32     1.1.1.1         1279            1278
*  >e 10.3.3.3/32     20.0.2.2        1277            1277

 Route distinguisher: 200:1
 Total number of routes: 2

     Network          NextHop         OutLabel        InLabel

*  >i 10.2.2.2/32     1.1.1.1         1278            1279
*  >e 10.4.4.4/32     20.0.2.2        1276            1276
<ASBR1>
<ASBR3>dis bgp routing-table vpnv4 inlabel | begin Network
     Network          NextHop         OutLabel        InLabel

*  >e 10.1.1.1/32     20.0.2.1        1278            1279
*  >i 10.3.3.3/32     6.6.6.6         1278            1277
```

图 5 私网的 LSP 状态结果

Fig. 5 LSP status results of private network

以 10.1.1.1 为例，阐述其私网的 VPN-LSP。图 5 中显示 10.1.1.1 的 outlable 为 1 279，其 nexthop 为 1.1.1.1，即说明该标签为 PE_1 分配给 $ASBR_1$ 使用；inlable 为 1 278，为 $ASBR_1$ 分配给 $ASBR_3$ 的标签。在 $ASBR_3$ 上的状态显示，outlable 为 1 278，inlable 为 1 279，其中 1 278 是 $ASBR_1$ 分配给 $ASBR_3$ 的，而 1 279 是 $ASBR_3$ 设备分配给 PE_2 的，图 5 所示的状态结果表明在域内与域间的私网 VPN-LSP 已成功建立。

4)私网用户数据互访测试

为验证在域内与域间能承载 U_1 与 U_3、U_2 与 U_4 的私网数据，在 PE_1 上执行 dis ip routing-table vpn-instance vpn1 | exclude Dir 与 dis ip routing-table vpn-instance vpn2 | exclude Dir，以及执行 ping -vpn-instance vpn1 -c 1 -a 10.1.1.1 10.3.3.3 与 ping -vpn-instance vpn2 -c 1 -a 10.2.2.2 10.4.4.4 指令，得到的结果如图 6 所示。图 6 中的结果显示，PE_1 的 vpn1 与 vpn2 实例路由表中通过 BGP 路由协议已学习到分部 U_3 与 U_4 的私网路由数据，并通过 ping 指令测试，U_1 与 U_3、U_2 与 U_4 私网数据交互正常。

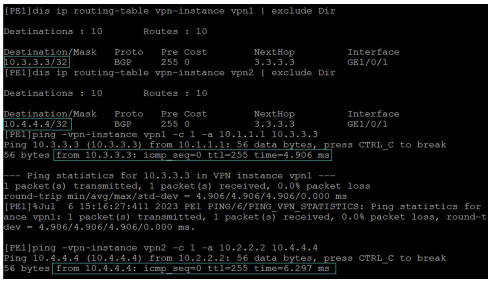

图 6 用户私网数据互访测试图

Fig. 6 User private network data mutual access test chart

3.2 性能分析

为了评估该方案的优越性，与传统的跨域最优方案-多跳的 EBGP 方式进行定性与定量的对比。

1)定性对比

定性对比包括可管理性、可靠性、资源利用率等 7 个方面。表 1 所列的 7 个维度的对比情况说明，采用虚拟化方式构建跨域私有网优于多跳的 EBGP 方式构建的跨域私有网，特别是在数据处理能力及设备的减负方面，优势明显。

表 1 两种方案的数据对比

Table 1 Data comparison of two schemes

对比项目	虚拟化方案	EBGP 方式
可管理性	多台设备虚拟化后，通过一台设备能管理其他设备，易管理	随着用户数增加，管理繁琐
可靠性	实现设备冗余，可靠性好	设备独立工作，会产生单点故障，可靠性差
资源利用率	虚拟后能实现资源统一调度与管理，利用率高	设备间独立工作，资源利用率低
需要维护标签类型	2 种	3 种
包转发率	高	低
扩展性	好	差
CPU 与内存使用率	只需处理 2 种类型标签，使用率低	需处理 3 种类型标签及 ACL 与路由策略，使用率高

2)定量对比

本文方案采集了传统方案与本方案自治系统边界设备 ASBR 的交换容量、路由条目、标签条目的数据，借助 EXCEL，导入了采集的数据，并用带数据标记的折线图生成了图表，对比情况如下文所述。

(1)交换容量

交换容量是衡量交换机能吞吐的最大数据量，是衡量交换机处理数据能力的重要指标。该方案采用网络设备虚拟化技术将位于自治系统边界的多台 ASBR 设备虚拟成一个资源池，让资源池中的设备共同负载分担其数据的处理，大大提升了设备的数据处理能力。图 7 给出了本文方案与传统的多跳 EBGP 方式采用 H3CS5820V2-54QS-GE 设备来构建跨域虚拟私有网的交换容量的数据对比结果，单台设备的交换容量是 11.52 Tbps。多跳 EBGP 方式构建的跨域虚拟私有网，因每台交换机是独立工作，增加交换机也不能实现交换机资源的整合，交换容量还是单台的处理能力。而本文方案能将多台交换机资源统一调度与管理，随着交换机数量的增加，其交换机处理数据的交换容量也相应地增长，当资源池的交换机数量为 9 时，其交换容量能达到 103.68 Tbps，使自治系统边界设备的数据处理能力得到明显的提升。

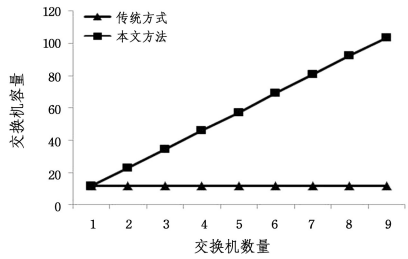

图 7 两种方案交换容量的对比

Fig. 7 Comparison of exchange capacity of two schemes

(2)路由条目

传统的多跳 *EBGP* 方式实现的跨域虚拟私有网，因私网数据的路由信息交互是在两台 *PE* 设备上直接交互，连接在两台 *PE* 内的所有设备都需要给 *PE* 的 loopback 地址分配标签，并将私网数据封装在标签中，才能实现私网数据穿越公网。而 MPLS 只能给域内的内部网关路由分配标签，域间是外部网关路由，MPLS 无法分配标签。为此，多跳 EBGP 方式修改了 MPLS 体系结构，采用 BGP 在域间的 ASBR 上交互 loopback 地址的路由，并为其分配标签，进而实现私网数据在域间的交互。但该方式使得位于自治系统边界的 *ASBR* 的 BGP 路由表中不但需维护用户的私网路由信息，还需维护 PE 的 loopback 地址的路由信息。而本文方案的 *ASBR* 的 BGP 路由表中不需要维护该路由，只需维护用户的私网路由信息。以该方案为例，采用本文方案 *ASBR* 的 BGP 路由表中只需维护 4 条路由条目，而采用传统的多跳 EBGP 方式需维护 6 条。设两自治系统的 *PE* 两两之间交互路由，不相互交叉交互路由，且后续增加到自治系统的 *PE* 所连接的每个用户有一条私网路由，一个 *PE* 连接两个用户，当两个自治系统的 *PE* 设备各达到 900 台时，采用本文方案只需要维护 3600 条路由条目，而采用传统的多跳 EBGP 方式需要维护 5400 条路由条目，如图 8 所示。图 8 中的状态结果表明，随着两自治系统内 *PE* 设备数量的不断增加，本文方案比传统的多跳 EBGP 方式需维护的路由条目数量少很多，减轻了 *ASBR* 的负载。

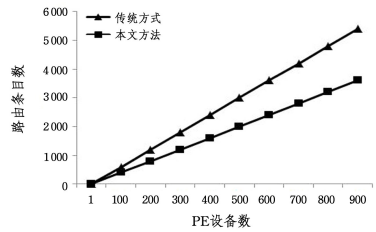

图 8 两种方案路由条目数的对比

Fig. 8 Comparison of the number of routing entries of two schemes

(3)标签条目

采用本文方案构建的跨域私有网，*ASBR* 只需维护 MPLS 分配的公网标签及 MP-BGP 分配的私网标签。而传统的多跳 EBGP 方式 *ASBR* 还需维护 BGP 给 PE 的 loopback 在域间分配的公网标签，且为了给 loopback 地址在域间分配公网标签，在 *ASBR* 上需给每个 loopback 部署访问控制列表来匹配 loopback 地址，并通过路由策略为 loopback 地址分配标签，两个自治系统的 *PE* 数增多，使得 *ASBR* 的 CPU 与内存的使用率加大，增加了 *ASBR* 设备的负担。图 9 给出

了本文方案与传统的多跳 EBGP 方式维护的标签数的对比，两自治系统各一台 *PE* 时，采用本文方案 *ASBR* 维护的标签数为 2，而采用传统的多跳 EBGP 方式需维护 4 条，随着两自治系统 *PE* 数的增加，两种方案维护的标签数的差距越明显。当各自治系统的 *PE* 数达 900 时，采用本文方案需要维护的标签数为 1800，而传统的多跳 EBGP 方式达到了 3600。对比数据说明本文方案给 *ASBR* 设备需处理与维护的标签数大大减少，减轻了 *ASBR* 设备的负担，优于传统的多跳 EBGP 方式。

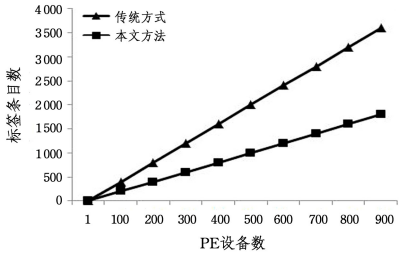

图 9 两种方案的标签条目数的对比

Fig. 9 Comparison of the number of label entries of two schemes

结束语 针对 MPLS 与 BGP 构建的跨域虚拟私有网接入的用户数增多，导致自治系统边界设备负载过重，影响用户数据的正常交互，本文提出了虚拟化的设计方案，构建网络模型，部署与实施虚拟化的设计方案，为解决该设计理念自治系统边界设备不能创建逻辑接口导致方案无法实施的问题，对该方案进行了优化。为评估方案的可行性与优越性，对虚拟化建立、公网的 LSP、私网的 LSP 等进行了测试与验证，验证与测试结果表明了方案的可行性。同时，将虚拟化构建的跨域私有网与传统的 EBGP 方式构建的跨域私有网在交换容量、路由条目、标签条目的 3 个方面，以及表 1 所列的 7 个方面，共计 10 个维度进行了对比。对比的结果表明，该方案在 *ASBR* 的数据处理能力及给 *ASBR* 减负方面，明显优于传统的方案。因学校网络实验室条件有限，无法测试方案的延时、并发数等，后续准备购买相应的设备，测试两种方案的延时与并发数等数据。此外，后续将研究软件定义网络(Software Defined Network，SDN)在跨域虚拟私有网中的应用。

参 考 文 献

[1] MA P Y, YANG G M, MAO D F, et al. Realization of intelligent routing based on SRv6+MPLS dual forwarding plane [J]. Optical Communication Research, 2022(1): 67-70.

[2] CHEN F Q, DOU J, ZHANG D. VRF configuration design and simulation based on BGP/MPLS VPN [J]. Journal of Chengdu University of Information Technology, 2020, 35(4): 378-381.

[3] LIN D S. Dual-machine hot standby function design of VPN gateway of industrial control system [J]. Computing Technology and Automation, 2020, 39(1): 74-78.

[4] WEN T, ZHANG Q B, AN W T. Construction of cross-regional technology video conference network based on VPN technology [J]. Journal of Liaoning University of Technology (Natural Science Edition), 2019, 39(3): 164-168.

[5] XU Y Y, SU XU Z, LIU Y Q, et al. Remote monitoring system for slub yarn production based on VPN technology [J]. Manufacturing Automation, 2022, 44(2): 16-19.

[6] SHENG W S, ZHOU C, SUN Y W. The application of MPLS VPN in enterprise networks [J]. Computer Technology and Development, 2020, 30(11): 117-122.

[7] SUN G Y, JIAO J. Simulation design of routing filtering based on multiple autonomous systems [J]. Journal of Capital Normal University (Natural Science Edition), 2022, 43(2): 20-28.

[8] SUN G Y, JIA Y X. Multi-Autonomous System Routing Simulation Based on GNS3 [J]. Laboratory Research and Exploration, 2019, 38(4): 123-128.

[9] LI Y F. Simulation design of enterprise cross-domain networking based on BGP MPLS VPN[J]. Laboratory Research and Exploration, 2021, 40(3): 121-128.

[10] DENG C R. Analysis on the principle of cross-domain VPN networking for power dispatching data network [J]. Journal of Shandong Electric Power College, 2022, 25(2): 20-24.

[11] SONG G J, HU C, ZHOU F. MPLS-VPN architecture optimization based on layered PE technology[J]. Computer Engineering, 2017, 43(6): 66-72.

[12] TAO Z Y, ZHANG J, YANG W D, et al. Research on performance optimization of edge devices based on double-layer virtualization idea [J]. Computer Science, 2021, 48(11): 372-377.

[13] SUN C. Ship navigation information service platform based on cloud computing virtualization technology [J]. Ship Science and Technology, 2022, 44(6): 141-144.

[14] HU Z Y. Design of intelligent ship automation information service platform based on cloud computing virtualization technology [J]. Ship Science and Technology, 2021, 43(22): 151-153.

[15] HUANG S P, XIE J, KAN H Y. Research on the Application of IRF Virtualization Technology in the Network[J]. Experimental Technology and Management, 2014, 31(11): 124-126.

[16] BAO L L, TANG H S, JIANG S Y, et al. Research on the design of meteorological network based on IRF2 and LACP MAD[J]. Computer Applications and Software, 2019, 36(1): 37-141.

[17] WANG J W, ZHANG X L, LI Q, et al. Research progress of network function virtualization technology[J]. Chinese Journal of Computers, 2019, 42(2): 185-206.

[18] SUN T, ZHAG J X. A review of virtualization technology research in network experimental beds [J]. Journal of Inner Mongolia University (Natural Science Edition), 2018, 49(5): 554-560.

TAO Zhiyong, born in 1980, master, associate professor, is a member of China Computer Federation. His main research interests include network communication and cloud computing.

ZHANG Jin, born in 1979, Ph.D, professor, Ph.D supervisor, is a member of China Computer Federation. His main research interests include network communication, cloud computing, and software engineering.

(责任编辑：喻藜)